CPP
Changing Professional Practice

Changing Professional Practice

Theory and Practice of Clinical Guidelines Implementation

Edited by Thorkil Thorsen & Marjukka Mäkelä

dsi INSTITUT FOR SUNDHEDSVÆSEN
Danish Institute for Health Services Research and Development

DSI rapport 99.05

CPP
Changing Professional Practice

DSI rapport 99.05

Changing Professional Practice

Theory and Practice of Clinical Guidelines Implementation

Health care professionals may have difficulties in keeping up-to-date with the overwhelming and fast growing volume of new scientific evidence for good clinical practice. Often there is an unacceptable time lag between the time of having a valid research basis for a specific procedure and the point when the procedure has been adopted in practice. This may lead to suboptimal care.

Clinical guidelines are regarded as an appropriate vehicle for overcoming this. They can assist health care providers in grasping the new evidence and bring it into daily clinical routines. But the production and dissemination of clinical guidelines does not automatically improve clinical practice. For instance, some types of guidelines and some dissemination strategies seem more effective than others; the effectiveness can differ across settings; the providers' willingness to use a guideline may depend on who produced it etc. There is a broad spectrum of circumstances that should be taken into account.

This book takes the reader through issues related to the introduction and use of clinical guidelines. In a multidisciplinary approach, it deals with concepts, methods, and theories relevant for studying the barriers and facilitators for the adoption of a guideline, the implementation process, the outcome, the costs etc.

Topics covered include:

- *A framework for guidelines implementation studies*
- *Design and statistical issues in implementation research*
- *Economic evaluation*
- *Qualitative approaches and methods*
- *Identifying barriers and facilitators*
- *Planning and monitoring interventions*
- *Measuring attitudes towards clinical guidelines*
- *A review of studies on perceived barriers and facilitators to implementation*

The book gives valuable information and tools for clinical guideline implementators and implementation researchers as well as for health policy decision-makers, purchasers, and providers.

dsi INSTITUT FOR SUNDHEDSVÆSEN
Danish Institute for Health Services Research and Development

CPP
Changing Professional Practice

Changing Professional Practice

Theory and Practice of Clinical Guidelines Implementation

Edited by Thorkil Thorsen & Marjukka Mäkelä

INSTITUT FOR SUNDHEDSVÆSEN
Danish Institute for Health Services Research and Development

DSI Rapport 99.05

DSI • Danish Institute for Health Services Research and Development

DSI • Danish Institute for Health Services Research and Development is an independent institution set up in 1975 by the state, The Association of Danish County Councils and Copenhagen and Frederiksberg municipalities. The institute's aim is to provide an improved basis for the work of the planning and management authorities within the health services.

DSI•Danish Institute for
Health Services
Research and Development
P.O. Box 2595
Dampfærgevej 22
DK-2100 Copenhagen Ø

Phone: +45 35 29 84 00
Fax: +45 35 29 84 99

Internet: http://www.dsi.dk
e-mail: dsi@dsi.dk

ISBN 87-7488-346-1
ISSN 0904-1737
DSI Rapport 99.05

Graphics production: Peter Dyrvig Grafisk Design / Formprint, Randers

Contents

Authors

Authors of this monograph

Marion K. Campbell, B.Sc. (Hons), M.Sc., Senior Statistician, Health Services Research Unit, University of Aberdeen, Aberdeen, Scotland.

Martin Eccles, MD, FRCP, FRCGP, MFPHM, Professor of Clinical Effectiveness, Centre for Health Services Research, University of Newcastle upon Tyne, Newcastle upon Tyne, England.

Marko Elovainio, Master of Soc. Sci., Ph.D. (Social Psychology), Senior Researcher, National Research and Development Center for Welfare and Health, Helsinki, Finland.

Jeremy M Grimshaw, M.B.Ch.B, Ph.D., M.R.C.G.P., Professor of Public Health and Programme Director, Health Services Research Unit, University of Aberdeen, Aberdeen, Scotland.

Richard Grol, Professor, Director of the Centre for Quality of Care Research of the Universities of Nijmegen and Limburg (W.O.K.), University of Nijmegen, Nijmegen, the Netherlands.

Lesley Hall, BA, Social Policy, Research Associate, Clinical Effectiveness Unit, Centre for Health Services Research, University of Newcastle upon Tyne, Newcastle upon Tyne, England.

Marlies Hulscher, Health Scientist, Ph.D., Senior Researcher, Centre for Quality of Care Research of the Universities of Nijmegen and Limburg (W.O.K.), University of of Nijmegen, Nijmegen, the Netherlands.

James Paul Kahan, Mathematical Social Psychologist, Ph.D., Program Manager, RAND Europe, Leiden, the Netherlands.

Miranda Laurant, Health Scientist, M.Sc., Junior Researcher, Centre for Quality of Care Research of the Universities of Nijmegen and Limburg (W.O.K.), University of Nijmegen, Nijmegen, the Netherlands.

Carl J. Lombard, M.Sc., Ph.D., Chief Statistician, Biostatistics Division, Centre for Epidemiological Research, Medical Research Council, South Africa.

Elaine McColl, BA, M.Sc., Social Scientist, Senior Research Associate, Centre for Health Services Research, University of Newcastle upon Tyne, Newcastle upon Tyne, England.

Emma McIntosh, Health Economist, Health Economics Research Unit, Dept. of Public Health, University of Aberdeen, Aberdeen, Scotland.

Jill Mollison, B.Sc. (Hons), Lecturer in Medical Statistics, Department of Public Health, University of Aberdeen, Aberdeen, Scotland.

Marjukka Mäkelä, MD, Ph.D., M.Sc. Head of Research and Development, National Research and Development Center for Welfare and Health, Helsinki, Finland. (Associate Scientist in Charge of CPP).

H. Claudia Pagliari, B.Sc, Ph.D., Lecturer in Psychology, Tayside Centre for General Practice, University of Dundee, Dundee, Scotland.

Nikki Rousseau, BA, Social Scientist, Research Associate, Centre for Health Services Research, University of Newcastle upon Tyne, Newcastle upon Tyne, England.

Nick Steen, Ph.D., Statistician, Senior Research Associate, Centre for Health Services Research, University of Newcastle upon Tyne, Newcastle upon Tyne, England.

Thorkil Thorsen, Sociologist, Ph.D., Acting Director of Research, DSI•Danish Institute for Health Services Research and Development, Copenhagen, Denmark. (Scientist in Charge and Co-ordinator of CPP).

Michel Wensing, Research Fellow, Ph.D., Centre for Quality of Care Research of the Universities of Nijmegen and Limburg (W.O.K.), University of Nijmegen, Nijmegen, the Netherlands.

Preface

Back in 1995 some European health services researchers under the lead of Finn Børlum Kristensen, Danish Hospital Institute (DSI), got together to plan a common project on clinical guideline implementation. The background for this was the often seen delay between the scientific documentation of best clinical practice on the one hand, and the uptake of this evidence into daily health care provider routines on the other. Implementation of well-grounded clinical practice guidelines was considered a useful way of speeding up this process – and of C*hanging Professional Practice*, which became the short title of the project (in daily speech and writing shortened to CPP).

A group of researchers from Finland, UK, Italy, the Netherlands, Norway, and Denmark had the application for a Concerted Action approved by the European Commisson's BIOMED-2 programme (contract no. BMH4-CT96-0697); and the project was launched in March 1996 with the title *Bridging the gap between science and practice: How to change health care provider behaviour through implementing clinical practice guidelines.* The project has been coordinated by DSI, now called Danish Institute for Health Services Research and Development.

Concerted Actions are co-ordination of research. This means that the national partner projects included in the overall project are to get their financial support from national funds and institutions. Unfortunately, this was not without problems for most of the partners. This resulted in the cancelling of one project and delays of others. The intended final systematic review of the partner project outcomes, therefore, could not be made within the stipulated project period.

But the project, which was due by the end of February 1999, has produced other deliverables according to the contract. This monograph is one of the main products. Its background and outline is described in the introductory Chapter 1.

Besides this monograph there have been other deliverables from the CPP-project, e.g. workshops, newsletters, lectures, instruments for measurement and calculations, and articles for peer reviewed journals. It is

our firm belief that the project has met most of the objectives possible and intended at the outset. Thus, much value has been produced within – i.e. for and by the participants of – the Concerted Action; and through the dissemination of the deliverables there is a community added value as CPP-outsiders have the opportunity of getting acquainted with and learning from the experience and results of the CPP.

We hope that this book can inspire and direct health policy decision-makers, purchasers, providers, clinical guidelines implementators as well as implementation researchers in their considerations of how to bridge the gap between science and clinical practice.

Acknowledgements

This book on clinical guideline implementation has been written by some of the researchers who participated directly or collaborated with our European Concerted Action "Changing Professional Practice". In addition to these authors, other participants and colleagues made important contributions through discussions of important issues, critical reviews of chapter drafts, provision of empirical data and experience from their own national projects. This book would have been impossible without their valuable contributions.

Therefore, on behalf of the authors and of the CPP-project, we want to give special thanks to the following members of the partner projects or working groups: Timo Sinervo, Jukkapekka Jousimaa, Pekka Rissanen, Alessandro Liberati, Signe Flottorp, Andy Oxman, Anne Walker, David Parkin, Eduardo Briones, Ignacio Marin, Lone Bilde, and Christina Holm-Petersen.

T. Thorsen & M. Mäkelä

Chapter 1

Introduction

Bridging the gap between science and practice: How to change health care provider behaviour through implementing clinical practice guidelines

Thorkil Thorsen and Marjukka Mäkelä

Clinical guidelines have quickly, and often with meagre justification, gained popularity in medical practice. They are usually defined as "systematically developed statements to assist practitioner and patient decisions about appropriate health care" (Institute of Medicine 1992). Guidelines are produced in various formats in most European countries and also by international professional organisations. Several countries (at least Finland, France, the Netherlands, Scotland, and the USA) have established programmes for guideline production during the 1990s.

This big interest in clinical guidelines may be seen as the answer to at least three circumstances. First, there is a growing awareness of large variations in clinical practice and other indicators of suboptimal health care. Second, the growing costs of health services call for a stricter and more rational health professional behaviour that may cut costs overall. And third, health professionals have difficulties in keeping up-to-date with the overwhelming and fast growing volume of new scientific evidence for good clinical practice.

Clinical guidelines are regarded as a vehicle for assisting health care providers in grasping the new evidence and bring it into daily clinical routines (this is what is alluded to in the title 'Bridging the gap between science and practice'), for improving practice, and for diminishing costs. Of course, clinical guidelines are not the only and sole solution to the problems mentioned.

Changing health care provider behaviour is only needed when the present provision of health care could be improved. But change in itself is no guarantee that things will get better. All change is not improvement. But all improvement is change.

Clinical guidelines may not always change or affect health services delivery once they are there. Several studies indicate that attitudes towards using guidelines may be positive, that providers know of a clinical guideline, but that the same providers do not use it in daily practice.

Our understanding of guideline use and effects has been patchy. To really bridge the gap between science and practice, and to take the theoretical discussions and developments to a higher level, a project pulling together European knowledge and experience was essential. Toward these aims, Changing Professional Practice (CPP) was established as a Concerted Action between eight European countries.

Our core issue is finding the means of facilitating better transfer of research findings into clinical practice using guidelines. This is a challenge to both health services research and health policy makers. Therefore it is essential that our results are available not only for the wider scientific community but also, and perhaps more importantly, for those who see guidelines as a useful tool in their search for better health services. That is the reason for publishing this monograph, which is one of the main products of the Concerted Action.

European research groups working with guideline implementation were brought together in the CPP project. Each group intended to study the implementation of at least one guideline in their own country. These empirical projects have provided material for the creation of a methodological framework. Guideline topics and environments for their application varied widely. Within CPP, we have studied the experiences of general practitioners in Finnish Lapland as well as those of thoracic surgeons in northern Italy. The applied guidelines have been either nationally or regionally developed. They have been published as traditional articles, pocket guides, on CD-ROM, or in Internet. Their lengths have varied from a dozen lines to booklets, and they have included text, hypertext, flow charts, photographs, and economic calculations. The providers using the studied guidelines have mostly been physicians, while some of the guidelines have been used by multiprofessional teams. Guidelines have been applied within primary, secondary, and tertiary care. Descriptions of the partner projects can be found in the Appendix.

Originally, we started out by finding controlled intervention studies that could contribute to a common analysis of what happens during

guideline implementation. The leading scientists of these studies jointly drafted a plan to collect selected data from their respective studies in similar format. The results of these studies were planned to provide data for a systematic review of the effectiveness of various guideline implementation strategies.

Guideline implementation studies are technically demanding. Study schedules have been changing over time, and by the end of the CPP in March 1999 only two projects were able to provide even preliminary data for joint analysis. The systematic review will therefore need to wait for some time. However, the methods created within CPP will make it possible to collect and systematically analyse results from not only the original CPP studies but also others conducted outside this umbrella.

In addition to the empirical studies several of the project partners in the Concerted Action also engaged in crossnational, multidisciplinary groups that have been responsible for several defined parts of the project, be they theoretical or methodological. The chapters of this volume reflect much of this work. Some of the groups have for the writing of their chapters invited CPP-outsiders to join as co-authors. (All authors of this volume are listed on pages 9-10).

In our terminology, "guideline" refers to a collection of statements ("recommendations") about a defined health problem, rather than to a single statement. All partner projects within CPP have used guidelines with multiple recommendations, typically covering several aspects of the diagnosis, treatment, and follow-up of the target disease or condition. Changing a set of actions, rather than a single action, requires time and effort.

Thus, "guidelines" should be distinguished from neighbouring concepts and quality assurance tools. For instance guidelines differ from *standards of quality* that are defined as "authoritative statements of (a) minimum levels of acceptable performance or results, (b) excellent levels of performance or results, or (c) the range of acceptable performance or results" (Grimshaw & Russell 1993). Eddy argues that "standards define appropriate care and should be followed in all circumstances with no flexibility for the clinician", which is not the case for guidelines which should be followed in most circumstances but allow for some flexibility (Eddy 1990).

Some words on the concept "implementation" are appropriate here. Ordinarily implementation means putting something (e.g. a plan or an innovation) into use. But in the clinical guideline jargon it has become commonplace to distinguish three ways of introducing or spreading clinical guidelines: diffusion, dissemination, and implementation. Lomas (1993) describes diffusion as a passive concept. Diffusion of e.g. infor-

mation on new medical evidence is not targeted, it is unplanned and uncontrolled, and only highly motivated potential recipients will seek it themselves. Dissemination is a more active process including launching of targeted and tailored information for the intended audience. Implementation will do more than only increasing the awareness of the target audience. "Implementation involves identifying and assisting in overcoming the barriers to the use of the knowledge obtained from a tailored message. It is a more active process still, which uses not only the message itself, but also organizational and behavioral tools that are sensitive to constraints and opportunities of identified physicians in identified settings." (*ibid.*).

It is this meaning of the term "implementation" that will prevail in this monograph when we speak about implementing clinical guidelines.

Studying clinical guideline implementation procedures, problems, costs and results is indeed a multidisciplinary enterprise. Our experience has made this distinctly clear. And this volume – organised into three main sections – is a reflection of this, containing chapters on different aspects of guideline implementation, and different approaches for studying how or why (or why not) a specific guideline is adopted by the clinicians and others who are to use it.

Section A lines up the battlefield. We have used the classification produced by the Cochrane Collaboration on Effective Practice and Organisation of Care Group (EPOC), with their permission, as a starting point for our work and for drawing up a framework for guidelines implementation studies. *Chapter 2* defines many different aspects and concepts closely related to guideline implementation research. It thus provides a common language. At the same time the chapter serves as a checklist for practical applications, e.g. for selecting aims, target groups, interventions, outcome measures.

The three chapters of *Section B* address methodological issues in guidelines implementation evaluation. *Chapter 3* advocates for the use of pragmatic in stead of explanatory studies of implementation strategies, which means that the effectiveness of an intervention should be measured in routine practice including 'normal' contextual and effect modifying factors, not under ideal conditions. The chapter discusses aspects of measurement (when to measure before and after, how to collect data), strengths and weaknesses of different study designs etc. Special attention is devoted to statistical issues and sample size calculation of cluster randomised trials.

Economic considerations become more and more an issue in the health care planning and prioritising. Economic evaluations of guideline implementation projects, however, have been rare so far. The CPP

workshops with participants from different countries and different disciplines and held during the Concerted Action indicated that one of the reasons for this might be the perceived complexity of identifying, measuring and valuing the costs and benefits of guideline production, implementation and use. *Chapter 4* discusses issues of costing and benefit assessment and techniques of economic evaluation of clinical guideline implementation strategies. A 'balance sheet approach' is presented which in a relative simple way identifies who bears the cost and who will have the benefits from any change. It is expected that this framework for economic evaluation will lead on to narrower cost-effectiveness or cost-utility analyses.

Throughout the CPP Concerted Action an urgent need for applying qualitative methods in guidelines implementation studies was identified again and again. Quantitative methods can describe 'how much', 'how many', 'which', 'when'. But when questions about 'how', 'what' 'why' and 'why not' are to be answered we must turn to qualitative data collection and analysis methods. Without the insights derived through such methods we will hardly be able to understand why some implementation strategies worked in some settings and in some professionals but not in others, or to find out whether it is the guideline itself or the implementation process that might be the problem. *Chapter 5* elaborates upon these issues, discusses different types of qualitative methods and their strong and weaker points, and also discusses the relationship and possible symbiosis of qualitative and quantitative studies.

Section C of this volume concentrates on instruments that are useful for planning the implementation of clinical guidelines. *Chapter 6* focuses on the different non-experimental methods that can be used for collecting data on and identifying barriers and facilitators to the implementation process and to the use of a clinical guideline. A pre-intervention measurement of these factors is useful for tailoring the implementation strategy; a measurement during or after the intervention helps understand (differing) results. As the experience is scarce, this chapter is more an overview of seemingly relevant methods than it is an evidence-based description of strengths and weaknesses of each.

Chapter 7 argues for the necessity of a thorough and minute planning, monitoring and description of interventions and strategies to implement guidelines in health care. This is relevant not only in relation to an individual project, but it can expand the possibilities of understanding the implementation processes and results also in a comparative perspective – across projects. For instance, the use of local opinion leaders to improve health professional practice has shown mixed results across projects (Thomson *et al.* 1999). But 'local opinion leader' as an intervention stra-

tegy may have different implications. Without detailed descriptions of how the opinion leaders were chosen, how they acted during the implementation phase etc., we will not be able to explain why some opinion leader-based interventions succeeded and others did not. Therefore, it is important to look inside the 'black box' of the intervention.

One of the prominent barriers to the adoption and routine use of clinical guidelines by the health care providers seem to be their general attitudes towards guidelines, or it might be their scepticism regarding the reliability of the guideline. *Chapter 8* describes the development and validation of a questionnaire on attitudinal aspects to guidelines implementation. The analysis revealed relationships between subscales of the questionnaire and the self-reported use of the guidelines. Such information is valuable for developing future guidelines.

Some of the barriers and facilitators that impede, resp. enhance the use of clinical guidelines are directly observable and easily measurable. But some of them, like the attitudes, are 'only' *perceived*, i.e. rooted in people's subjective beliefs. They are, nevertheless, real in the sense that they interfere with a smooth uptake and use of a clinical guideline. *Chapter 9* is a bibliography of studies that have investigated the *perceived* barriers and facilitators for guideline implementation and use. This 'library' of studies is categorised, so that the reader can look for references to studies with a certain focus, target groups, data collection methods etc. Most of the studies have used structured survey/interview methods, but other methods are represented as well.

As mentioned above, this book is one of the products of the CPP Concerted Action. In addition, both articles and monographs with more detail will be published in various fora for the scientific community. Many products from the Concerted Action are available for the user at the DSI website http://www.dsi.dk/ where the full-text version of this monograph also will be reachable. We hope that this book is useful for future researchers and active proponents of guideline implementation.

References

Eddy DM. Practice policies – what are they? *JAMA* 1990; 263: 877-80.

Grimshaw J, Russell I. Achieving health gain through clinical guidelines. I: Developing scientifically valid guidelines. *Quality in Health Care* 1993; 2: 243-48.

Grimshaw JM, Russell IT. Achieving health gain through guidelines II: Ensuring guidelines change medical practice. *Quality in Health Care* 1994; 3: 45-52.

Institute of Medicine. *Guidelines for clinical practice: From development to use.* Washington DC: National Academic Press 1992.

Lomas J. Diffusion, Dissemination, and Implementation: *Who Should Do What? Annals of the New York Academy of Sciences* 1993; 703: 226-35.

Thomson MA, Oxman AD, Haynes RB, Davis DA, Freemantle N, Harvey EL. Local opinion leaders to improve health professional practice and health care outcomes. (Cochrane Review). In: *The Cochrane Library*, Issue 1, 1999. Oxford: Update Software.

Section A

Framework

Chapter 2

A framework for guidelines implementation studies

Marjukka Mäkelä and Thorkil Thorsen

Aims of the chapter

The main purpose of this framework is to facilitate guideline implementation by providing a common language and a checklist for practical applications. In order to increase our knowledge or to change actual practice most efficiently, plans for and reports on systematic studies or practical implementation projects on clinical guidelines should be made in a structured fashion. This checklist is meant to assist in defining the aims, the target groups, the methods, the interventions, and the outcomes of guideline implementation projects clearly and in a manner that also allows for reliable comparison between various projects.

Introduction

This framework attempts to list the dimensions and attributes of guidelines and their implementation, both subjective and objective, that may affect the process of professional practice change. As a descriptive instead of explanatory framework, it does not attempt to evaluate the importance of these dimensions as facilitators or barriers to guideline implementation. It can be used by decision-makers who want to improve practices, by reviewers, and by researchers. The framework outlines a uniform terminology, providing a field-tested answer to "What do we mean, when we talk about X, an aspect of guideline implementation".

The purpose of constructing and using a common framework is to improve the accessibility, usefulness and – most importantly – comparability of the data that we have from the different studies on changing professional practice in health care. In developing the framework, we have tried to avoid forcing data into a theory. Instead, we have been able to develop the framework in an iterative way, starting from previous work in the area, using data available from our various projects within the Concerted Action of the Changing Professional Practice (CPP) project, and applying an earlier version of the framework to partner project descriptions. A special CPP group has worked on developing this framework for guideline implementation. The members of the group are listed at the end of this chapter.

The selection of items for this framework is based on their relevance to the implementation of guidelines. This framework is not meant to be used as a checklist for guideline planning or production. There are checklists for ensuring various aspects of guideline validity (e.g. Grimshaw *et al.* 1995) and others that look at the appropriateness of guidelines production programmes (Lohr 1998). These elaborate on issues listed here as relating to the guideline itself, and ideally they can match this framework, bringing in more detail. This CPP checklist can be used as a part of a systematic framework for evaluating guideline implementation studies.

In this chapter, the framework developed in dialogue with the CPP partner projects is described under six main headings. Our focus is on trials, although we recognise the necessity of using other study designs as well. Included are the areas that have seemed most crucial to our various implementation studies. Several parts of the framework have been developed in close connection to other CPP subgroups, also drawing from general theory in the relevant areas (for example, economics or study design). Another starting point has been to apply the definitions and concepts already elaborated upon by others as much as possible, instead of constructing new ones.

The most important group working with these issues previously has been the Cochrane Collaboration on Effective Professional Practice (CCEPP), today called the Cochrane Effective Practice and Organisation of Care Group, or EPOC. The structure of the framework proposed here has got much inspiration from the CCEPP Data Collection Checklist from March 1996, and some of the items are close or identical to theirs. Using the CPP data, we have reclassified many of these items and developed new ones. Consequently, we expect that this framework will need to be reviewed as the art of studying guidelines implementation develops with experience and new methodologies.

For a quick overview of the framework, the main categories are outlined below. In the succeeding sections each category is described more thoroughly, including subcategories and definitions. The lowest level categories are in italics.

Clinical practice guidelines
- Level of development
- Organisation behind guideline
- Intended sector of application
- Representation in guideline group or guideline development process
- Development method
- Process of guideline authorisation and adaptation
- Aim of the clinical guideline
- Updating the guideline
- Production of guideline as part of a guidelines programme

Interventions
- Implementation strategies
- Format
- Content
- Sender/deliverer
- Type of targeted action

Participants and settings
- Characteristics of providers
- Characteristics of participating patients

Processes and Outcomes
- Health professionals
- Patients
- Organisational processes and outcomes
- Economic processes and outcomes

Barriers and facilitators
- Professionals
- Patients
- Environment

Methods
- Study design
- Unit of analysis
- The number included in the study

1. Clinical practice guidelines

Clinical practice guidelines vary in scope and length. Within a guideline, there are usually a number of *specific statements*, each supporting one decision. 'Guideline' may thus refer to a single specific statement (such as choice of medication in a given clinical situation) or, more often, to a collection of statements about the prevention, diagnosis and treatment of a health problem (such as treatment of asthma or breast cancer). In this framework, *guideline* refers to a collection of statements about a defined topic. Guidelines are not laws; even if it is usually wise to follow them in most circumstances, they should also allow for individual flexibility.

Medical *review criteria* are 'systematically developed statements that can be used to assess the appropriateness of specific health care decisions, services, and outcomes' (Institute of Medicine 1992). Such criteria help to evaluate to what extent a *specific statement* within a guideline has been applied (e.g. what percent of asthma patients received the recommended medication). *Standards of quality* are authoritative statements of minimum levels or a range of acceptable performance or results, or they can define excellent levels of performance or results (Institute of Medicine 1992). Guidelines may in some environments contain review criteria for which standards have been defined.

Clinical guidelines can be described in terms of the level at which they were developed and the sector for which they were meant. In addition, it's important to know who participated in the development, how they were developed, authorised, adapted, and updated and what their aims and targets are.

Many aspects of guidelines are not listed here, either due to difficulties in classification (for example, guideline complexity) or because their relevance to implementation is not known. Active development of the knowledge base in this area is done in various countries, and also in a BIOMED collaborative project aiming at the development of an appraisal instrument for clinical guidelines (Littlejohns 1999, personal communication).

1.1. Level of development

Clinical guidelines can be developed at different levels, e.g. at the local or the central level. The level of development may carry great significance, as e.g. familiarity with the local organisation of health care, or the authority of the guideline developers, may influence the (perceived) appropriateness of the guideline and the willingness of the targeted audience to use it.

1.1.1. Local:
A local guideline development group consists mainly of persons from a local professional culture and/or a small geographical area; these persons collaborate on a daily basis in providing care for patients targeted by the guideline. The group may include or consult methodological or content experts from outside the local area.

1.1.2. Regional:
A regional guideline development group includes persons from a wider geographical area than the local one, typically from several local units. All targeted units need not have representation. The group may include experts as above.

1.1.3. National:
A national guideline development group consists of nationally or centrally chosen persons, often representing various interest groups (e.g. speciality societies or central health authorities). It may or may not include clinicians who are supposed to be the users of the guideline.

1.1.4. Supranational:
A supranational guideline development group consists of representatives from several countries and/or international organisations and agencies (e.g. international professional organisations or WHO).

1.2. Organisation behind guideline (source)

The organisation responsible for producing the guideline, or guideline source, carries a significance in its implementation. The willingness to accept and apply guidelines can depend at least partly on the perceived motivation of the producer. Independent, professional organisations seem to carry least burden of doubt as guidelines producers.

1.2.1. Professional organisation

1.2.2. Administrative non-funding unit:
Ministries or units funded by them, regional governmental health units and other similar units that do not have the power of directing resources to health care providers.

1.2.3. Health care funding organisation:
Health insurance companies, governmental units with resource allocation power, etc.

1.2.4. Health care delivery unit:
Hospitals, health centres, etc.

1.2.5. Others, describe

1.3. Intended sector of application

A clinical guideline may be targeted exclusively at either primary care (general practitioners, health centres) or the hospital sector. Increasingly, guidelines are being produced for the entire health care process, during which the patient moves from primary to tertiary care and back. It is not always clear which sector a guideline originally was produced for, and some guidelines are being used outside their original target group.

1.3.1. Primary care:
This means care by general practitioners or other health care personnel that covers prevention, diagnosis, treatment and rehabilitation for common health problems. Patients usually have access to primary care without referral.

1.3.2. Secondary care:
Usually delivered by specialists in other disciplines than primary care (e.g. surgery, gynaecology, internal medicine), for either inpatients or at outpatient clinics. Access to secondary care may or may not require a referral from primary care.

1.3.3. Tertiary care:
Delivered in institutions to which referral usually is required from primary or secondary care; represents a higher level or specialisation than secondary care (e.g. plastic surgery, fertility treatments, endocrinology). Super-specialists, with narrower speciality definition and/or higher level of technology, may require a referral from within tertiary care but are included in this category.

1.4. Representation in guideline group or guideline development process

Guideline development may take place in various stages (producing the

draft, comment rounds, finalising and updating the guideline) and the involvement of relevant actors may need to be specified separately for each stage. If unspecified, participation means an active role in the formulation of a guideline at any stage. One may specify whether the participants have been appointed as official representatives of an organisation or selected in a more informal fashion.

1.4.1. Users involved:
The professional end users, i.e. the relevant health care professionals, have actively participated in the guideline development group or at some other stage in the guideline development process.

1.4.2. Experts involved:
Experts, i.e. persons with scientific or practical background from relevant fields (such as medical specialities, nursing sciences, social sciences, economics etc.) have participated in the guideline development group or process.

a. Experts on the clinical content
   i) Primary-care expert
   ii) Secondary-care expert on clinical speciality in question
   iii) Expert on other relevant clinical speciality
   iv) Allied health professions (nurse, midwife etc.), specify
b. Experts on other content
   i) Health economics
   ii) Social sciences (sociologist, psychologist etc.), specify
   iii) Other, specify
c. Experts on guideline development and methodology
   i) Informatics, library sciences
   ii) Clinical epidemiology
   iii) Statistics
   iv) Communication
   v) Other, specify

1.4.3. Patients involved:
Patients with the targeted disease(s) or condition(s), or their representatives, have participated in the guideline development group or process

a. Patients suffering from the disease or condition in question
b. Representatives of patient organisations

c. Representatives of relatives' organisations
d. Other, specify

1.4.4. Others involved:
Other interest groups, stakeholders, or organisations (e.g. administrators, funders, politicians) have been represented in the guideline development group or process.

Specify

1.5. Development method

Clinical guidelines are classified also according to how scientific evidence and data on clinical practice have been identified for and used in guideline development. Methods may be described in the guideline, its background document or in the documentation of the guidelines programme. The methods may differ for various statements within a guideline, and should be classified according to the main method used throughout the guideline.

1.5.1. Methods of identifying evidence:
Method(s) used for finding the scientific evidence for guideline statements, as described in the guideline or its background documentation.

a. Systematic review (clearly described method of finding evidence, results presented as a meta-analysis or in some other systematic manner)
b. Non-systematic review (methods of finding evidence not described or not repeatable)
c. Experience-based method, i.e. based on guideline developers' own professional experience or experience as patients
d. Not described or not known

1.5.2. Methods of identifying data on practice patterns:
Method(s) used in collecting data about actual practice patterns in the area(s) for which the guideline is targeted and/or used as a basis for the construction of the guideline. These can include structure, process or outcome data that are relevant to the guideline.

a. National statistics
b. Practice surveys of several clinical units, including units outside the guideline development group
c. Audits of guideline developers' own practice
d. Not described or not known

1.5.3. Methods of deriving the recommendations: Methods for establishing links between evidence/documentation and guidelines.

a. Explicit evidence-linked method: The guideline statements are clearly linked to relevant evidence. In addition, the level of evidence behind each statement is searched for and evaluated systematically and described transparently. This may include representing the strength of individual studies or their combined results using pre-defined categories. In addition, other methods (consensus etc.) may be used for recommendations for which evidence is contradictory, scarce or lacking. The explicit method may include evaluation and calculation of the potential benefits, risks and costs of different interventions. The guideline describes the basis for these calculations.
b. Evidence-linked method: An expert group develops the guideline through a review of scientific evidence relevant to the recommendations in the guideline. There is linkage between scientific documentation and the recommendations, but it is not explicit and/or the review is not systematic.
c. Formal consensus method: Here the basis for guideline development is a consensus conference, nominal group technique, Delphi technique or the like. Rules for the analytical process exist, but explicit linkages between scientific documentation and recommendations are not defined. Therefore, the composition of the expert panel and their judgements are of final importance.
d. Informal group method: The development and decisions are made through open discussions in an expert group, but the criteria for decision-making are poorly defined, or only implicit. Scientific information is used in an unsystematic way; the final document does not record how consensus was reached, or to what degree the recommendations are based on scientific documentation vs. subjective judgement.
e. Not described.

1.6. Process of guideline authorisation and adaptation

Guidelines developed outside the setting where they will be used may not be acceptable for users unless an active adaptation process or authorisation procedure has taken place. Regionally developed guidelines may be further adapted locally. Similarly, nationally developed guidelines can be tailored for regional or local use.

1.6.1. Formal authorisation before or during introduction of guideline

1.6.2. Pre-testing before guideline introduction

1.6.3. Adaptation or local/regional elaboration before implementing the guideline in the local/regional clinical culture.

1.7. Aim of the clinical guideline

A guideline may have one or more specific aims. These can be stated explicitly for a single guideline or for a guidelines programme. The aims may also be specified separately for each of the major recommendations in the guideline. Within one guideline, several different types of changes can be proposed; and within a type of change, multiple actions can be suggested (e.g. modification of established management can include several different components, such as increasing one activity while reducing or discontinuing another one).

1.7.1. Purpose of recommendations

a. Appropriate clinical care
b. Reducing management variation (usually assumes variation is shown to exist)
c. Cost containment
d. Other, specify
e. Unspecified

1.7.2. Nature of desired change

a. Initiation of new management (e.g. introduction of a new technology/method)
b. Discontinuing introduction of new management
c. Preventing introduction of new management
d. Increase in established management
e. Reduction of established management
f. Cessation of established management

1.8. Updating the guideline

Guidelines are rapidly outdated, as new research results become available. Ideally, guidelines have a 'best before' date telling when the contents will be reviewed to incorporate new information.

1.8.1. Guideline is updated on a regular basis:
It is stated in the guideline when it is to be updated, or the guidelines programme of which the guideline is a part includes regular updates.

1.8.2. Guideline is updated irregularly:
It is stated in the guideline under which conditions it is to be updated.

1.8.3. Guideline is not updated:
The guideline is meant to be used unchanged, or there is no mention of when or under which conditions it is to be updated.

1.9. Production of guideline as part of a guidelines programme

Increasingly, guidelines are being produced in regional or national programmes that to some extent at least ensure the quality of production. The availability of guidelines and motivation to use them may increase when end users know there is a variety of guidelines available.

1.9.1. Produced in a permanent guidelines programme:
The guideline is one of an identifiable set of guidelines that share at least some aspects of methodology in their production, distribution and/or implementation.

1.9.2. Produced as one of several guidelines using similar methods:
There are several guidelines produced within a project or using similar methods, but the programme for guideline production is not continuous.

1.9.3. Produced as a single guideline:
No links to other guidelines

1.9.4. Not known.

2. Interventions

The introduction of clinical practice guidelines may be supported by active implementation strategies and methods. These should facilitate guideline use and help overcome barriers to their adoption in clinical settings. Parts 2.1 and 2.5 of the framework are based on the work from EPOC, with updates in part 2.1.3 and the corresponding part 4.3. Several Cochrane and other reviews on previous studies of health care interventions have been done using the EPOC (CCEPP) March 1996 checklist. Parts 2.2 to 2.4 are based on theoretical work presented in more detail in Chapter 7.

2.1. Implementation strategies

In most guideline implementation projects, several implementation strategies are combined to a functional program. It may be difficult to clearly separate the organisational part of the intervention from the part directed toward professionals, for example – the latter may be partly nested within the former. Strategies can also occur in a series, and a strategy may be an outcome of a guideline implementation project: A staff-oriented change (for example the introduction of teamwork or case management) may enhance patient participation. For easiest possible interpretation of intervention study results, a serious attempt at a clear description of the various aspects is highly recommendable.

2.1.1. Interventions orientated toward health professionals

a. Distribution of educational materials: Distribution of published or printed recommendations for clinical care, including clinical practice guidelines, audio-visual materials and electronic publications. The materials can be delivered personally or through mass mailings.
b. Conferences: Participation of health care providers in conferences, lectures, workshops or training sessions outside their own practice settings.
    i) Small-group conferences (active participation)
    ii) Big-group conferences (passive participation)
c. Local consensus processes: Inclusion of participating providers in discussion to ensure that they agree that the chosen clinical problem is important and the approach to manage the problem is appropriate.
d. Outreach visits: Use of a trained person who meets with providers in their practice settings to provide information. The information given may include feedback on the providers' performance.

e. Local opinion leaders: Intervention using providers nominated by their colleagues as 'educationally influential'. It should be explicitly stated how the opinion leaders were identified (by their colleagues) and how they were recruited.
f. Patient-mediated interventions: Any intervention aimed at changing the performance of health care providers where information was sought from or given directly to patients by others; e.g. direct mailings to patients, patient counselling delivered by others, materials given to patients or placed in waiting rooms.
g. Audit and feedback: Any information or summary of clinical performance in health care over a specified period of time. The information may be given in a written or oral format and it may also include recommendations for clinical action. Information on provider performance may have been obtained from medical records, computerised databases, observation, or from patients.
    i) Internal audit, i.e. audit performed by the providers themselves
    ii) External audit, i.e. the providers getting data on their performance from others.

        The following interventions should **not** be included in the audit and feedback:

        i) Provision of clinical information not directly reflec-ting provider performance collected by the investigators directly from patients, e.g. scores on a depression instrument
        ii) Feedback from individual patients' health record information in an alternate format (e.g. computerised).
h. Reminders: Any intervention, manual or computerised, that prompts the health care provider to perform a clinical action. The following interventions are included:
    i) Computerised decision support. (Use of an active knowledge system which uses two or more items of patient data to generate case-specific advice)
    ii) Concurrent reports. (Targeted at providers at the time of an encounter to remind them of desired actions for individual patients)
    iii) Intervisit reminders. (Targeted at providers between visits when there is evidence of suboptimal care for specific patients, e.g. when a test is abnormal and the appropriate follow-up is not found in the medical record)

iv) Enhanced lab report. (Lab report following abnormal result targeted at providers which includes additional information about specific follow-up recommendations)
v) Administrative support. (Follow-up appointment systems or stickers on charts)
vi) Implicit reminders. (Predictive values for abnormal test results without an explicit recommendation for action).

i. Tailored interventions: Use of personal interviewing, group discussion ('focus groups') or a survey of targeted providers to identify barriers to change and subsequent design of an intervention that addresses identified barriers.
j. Peer review
k. Combined strategies: Specify using the above classification of interventions, if applicable.
l. Other, specify

2.1.2. Financial interventions

a. Provider interventions
   i) Fee-for-service (Provider is paid a fixed amount for the number and type of services delivered)
   ii) Capitation (Provider gets paid a set amount per person in the target population per time unit for providing specific care)
   iii) Provider salaried service (Provider gets basic salary for providing specific care)
   iv) Provider incentives (Individual provider gets direct or indirect financial reward or benefit for doing specific action)
   v) Institution incentives (Institution or groups of providers get direct or indirect financial rewards or benefits for doing specific action)
   vi) Provider grant/allowance (Individual provider gets direct or indirect financial reward or benefit not tied to specific action)
   vii) Institution grant/allowance (Institution or groups of providers get direct or indirect financial reward or benefit not tied to specific action)
   viii) Provider penalty (Individual provider gets direct or indirect financial penalty for inappropriate action)
   ix) Institution penalty (Institution or groups of providers get direct or indirect financial penalty for inappropriate action)

x) Changes in formulary (Additions or removals from reimbursable available products)
xi) Other (specify)

b. Patient interventions
   i) Premium
   ii) Co-payment
   iii) User-fee
   iv) Patient incentives
   v) Patient grant/allowance
   vi) Patient penalty
   vii) Other, specify

2.1.3. Organisational interventions:
These may include changes in the physical structures of health care units, in medical record systems or in ownership.

a. Structural interventions
   i) Changes in the settings/site of service delivery (e.g. moving a family planning service from a hospital to a school)
   ii) Telemedicine (providing means of communication and case discussion between distant health professionals)
   iii) Changes in medical records systems (e.g. changing form paper to computerised records)
   iv) Other changes in arrangements for maintaining or retrieving information (e.g. patient tracking system)
   v) Other changes in physical structure, facilities and equipment
   vi) Changes in scope and/or nature of services (e.g. introducing day surgery)
   vii) Changes in presence and organisation of quality management mechanisms
   viii) Changes in ownership and/or affiliation status of hospitals and other facilities
   ix) Other structural changes in organisation excluding staff (specify)

b. Staff-oriented interventions
   i) Revision of professional roles (Changes in role contents among health professionals also known as 'professional substitution' or 'boundary encroachment'; e.g. nurse midwives providing obstetrical care or pharmacists providing drug counselling that was formerly provided by nurses and physicians)

ii) Multidisciplinary teams (Health professionals of different disciplines work together as a team to care for a patient or population)
iii) Case management (One professional takes responsibility for co-ordinating care given to one patient by several providers and/or units)
iv) Other integration of services (Follow-up mechanisms to co-ordinate a patient's care across organisational or unit boundaries; sometimes called 'seamless care')
v) Skill mix interventions (Changes in numbers, types or qualifications of staff)
vi) Interventions to improve provider satisfaction with the conditions of work or its material/psychic rewards (E.g. interventions to 'boost moral')
vii) Other, specify

c. Patient-oriented interventions
i) Interventions facilitating individual patient participation (E.g. decision support tools for patients)
ii) Interventions facilitating patient group participation (E.g. focus groups, patient panels)
iii) Other, specify

2.1.4. Regulatory interventions:
Any intervention that aims to change health service delivery or costs by regulation or law. These interventions may overlap with organisational and financial interventions, or one intervention may contain elements from several categories.

a. Changes in medical liability
b. Management of patient complaints
c. Accreditation
d. Licensure
e. Other (specify)

2.2. Format

The format of the intervention consists of a variety of factors, which often are linked together. Usually it is best to describe the format for each specific intervention strategy.

2.2.1. Timing: Scheduled and/or actual timing of the various parts of the intervention

a. Proximity: Time interval between delivering the intervention and clinical decision-making.
    i) Prospectively linked to patient care decision (During the decision, as for example computer reminders)
    ii) Retrospectively linked to patient care decision (Immediately after the consultation)
    iii) Not linked to patient care decision (Summary feedback report etc.)
b. Schedule: Number and duration of intervention events and intervals between these events.
    i) Number of intervention events
    ii) Duration of each similar intervention
    iii) Frequency (Time intervals between intervention events)

2.2.2. Media:
Type(s) of material(s) used in the delivery of the intervention

a. Oral
b. Written
c. Electronic
d. Other, describe
e. Combination, describe

2.2.3. Flexibility: Variation in the delivery of the intervention allowed

2.3. Content

The content of the provided information includes guideline-related information and data related to performance. This part of the framework overlaps slightly with part 2.1; as the interventions described in 2.1 have been used in various reviews, we find it useful to provide both sets of definitions here.

2.3.1. Type of information

a. General or background information
b. Specific information on guideline topic
c. General information on practice variation
d. Specific feedback on own performance
e. Other, specify

2.3.2. Presentation:
Mode of presenting data used in the intervention

a. Qualitative: descriptive
b. Quantitative: numerical, graphical

2.3.3. Comparability of information:
Possibility to compare own actions or performance with others.

a. Comparability with guideline content
b. Comparability with other standards
c. Comparability with other health care units/ practitioners

2.4. Sender/deliverer

The credibility of the intervention depends partly on the individual or the group delivering the intervention. Both the organisational (as discussed in 1.2.) and the individual sources can be experienced as having or lacking scientific credibility, other authority, regulatory power or various open or hidden agendas.

2.4.1. Person(s) delivering the intervention

a. Local expert, specify profession
b. Central expert, specify profession
c. Research worker, specify
d. Management representative
e. Computer system
f. Other, specify

2.4.2. Authority

a. Credibility (Level or knowledge, membership, etc.)
b. Attractiveness (Familiarity, manner of interaction, etc.)
c. Power (to give rewards or punishments, etc.)

2.5. Type of targeted action

Early in the time of guidelines, it was common to implement narrow, topic-specific interventions that targeted a very specific action. Today these are often embedded as smaller parts or single statements within a guideline. It is useful, however, to list the various types or units of targeted action as part of the rich description of guidelines implementation interventions.

2.5.1. Individual preventive services

2.5.2. Individual health promotion/advice

2.5.3. Preventive services at population level

2.5.4. Population health promotion

2.5.5. Diagnosis

a. laboratory
b. imaging
c. other, specify

2.5.6. Test ordering

2.5.7. Referrals

2.5.8. Procedures

2.5.9. Prescribing

2.5.10. Professional-patient communication

2.5.11. Record keeping

2.5.12. Discharge planning

2.5.13. Other resource use (specify)

2.5.14. Other (specify)

3. Participants and settings

3.1. Characteristics of providers

Several types of providers may be targeted by the same intervention either directly or indirectly. Some barriers and facilitators to implementation may be specific to profession, and sometimes we may want to tailor the interventions for various target groups separately. For these and other reasons, a detailed description of the participants and settings is necessary in intervention studies, and also for interventions without research ambitions it is useful to describe their target groups to be able

to think more clearly about the intervention itself. For example, auxiliary health personnel may start using guidelines that originally were directed at physicians; suitably edited versions of guidelines for this target group can then be prepared. Researchers and implementers also need to be alert to observe possible country-specific variations in settings, reimbursement and division of labour.

3.1.1. Profession

a. Physicians
b. Nursing staff
c. Pharmacists
d. Physiotherapists
e. Dentists
f. Psychologists
g. Mixed
h. Other providers (specify)

3.1.2. Clinical speciality

a. General/family practice
b. Internal medicine
c. Geriatrics
d. Surgery
e. Psychiatry
f. Paediatrics
g. Obstetrics & gynaecology
h. Laboratory medicine
i. Radiology
j. Other (specify)

3.1.3. Age and gender of providers

a. Age: mean and range, or age groups, by gender
b. Gender
    i) >90% females
    ii) 65%-90% females
    iii) more than 35% of both genders
    iv) 65%-90% males
    v) >90% males

3.1.4. Seniority

a. Training
   i) Student
   ii) Licensed practitioner
   iii) Specialist
   iv) Other
   v) Not applicable
b. Administrative activity
   i) None
   ii) Clinical and administrative duties
   iii) Administrative duties only
   iv) Other
   v) Not applicable

3.1.5. Reimbursement system

a. Predominantly private funding
b. Predominantly public funding

3.1.6. Setting of care

a. Sector
   i) Primary care
   ii) Secondary care
   iii) Tertiary care
b. Unit of care
   i) Individual professional
   ii) Team
   iii) Ward
   iv) Institution
c. Patient contact
   i) No (e.g. laboratory test)
   ii) Yes, continue:
      1. Patient's place of stay
         a. Short visit(s) to unit of care, no overnight stay
         b. Day care, no overnight stay
         c. Institutional stays at the institution
      2. Character of contact with the health-care setting
         a. Acute
         b. Elective

3. Duration of contact with the health-care setting
    a. Occasional, e.g. patients visiting the unit of care on a limited number of occasions (without an agreed-upon plan)
    b. Short term, i.e. patients with a planned short duration of continuous contact (days or weeks, not months)
    c. Long term, i.e. patients with a planned long duration of continuous contact (for months)

3.1.7. Main target group
    a. Individual providers
    b. Groups of providers with the same profession (size)
    c. Multiprofessional teams (size)

3.1.8. Social interaction: Qualitative description

3.1.9. Motivation for participation
    a. Voluntary or compulsory
    b. Financial incentives: yes/no
    c. Other motivation, describe

3.1.10. Country

3.2. Characteristics of participating patients

3.2.1. Clinical area: Specify (e.g. diabetes, clients at well-baby clinics etc.)

3.2.2. Other characteristics
    a. Age
    b. Gender
    c. Ethnicity
    d. Other, specify (e.g. social group)

4. Processes and Outcomes

The outcome measures reflecting changes achieved at different levels should be specified well in advance. An obvious main outcome would be a change in morbidity or mortality, or if such effects can't be mea-

sured, an increased compliance with the guideline. In addition to or instead of these, there may be multiple secondary outcomes observable in health professionals, patients, organisations or the economics of health care in the area of the guideline. Some measured effects (mainly knowledge, skills and attitudes) may but do not necessarily result in actual outcomes; these are proxy effects and should be treated as such. All indicators of effect should be measured in a reliable fashion.

4.1. Health professionals

4.1.1. Compliance with guideline, clinical action

4.1.2. Other action

4.1.3. Factors facilitating main outcomes

a. Knowledge
b. Attitudes
c. Skills
d. Satisfaction (with guideline, own performance, or work)

4.1.4. Other, specify

4.2. Patients

The main purpose of health care is to prevent decline in or improve the health of the target population. The main outcomes of guideline implementation projects should therefore be decreases in mortality, morbidity or other important clinical parameters measurable in patients targeted in the guideline.

4.2.1. Mortality

a. Mortality from the guideline-targeted disease or condition
b. Mortality from all diseases or conditions

4.2.2. Morbidity

a. Morbidity from the guideline-targeted disease or condition
b. Morbidity from all diseases or conditions

4.2.3 Clinical parameters (e.g. blood pressure, peak flow)

4.2.4 Indicators of service use (e.g. readmission, length of stay)

4.2.5 Compliance with medical advice:
Behaviour specifically targeted in the guideline (e.g. participating in screening, following a drug regimen)

4.2.6 Other behaviour:
Behaviour change not targeted in the guideline (e.g. coping, social support, change in patient organisation)

4.2.7 Factors facilitating main outcomes

a. Knowledge
b. Attitudes

4.2.8 Satisfaction

4.2.9 Quality of life

4.2.10 Other, specify

4.3. Organisational processes and outcomes

These include a variety of changes, which may or may not be designated in the guideline. The structure in this part is similar to that of part 2.3.1 in the framework; these differ in several details from the EPOC classification.

4.3.1. Structural changes

a. Changes in the settings/site of service delivery
b. Telemedicine
c. Changes in medical records systems
d. Other changes in arrangements for maintaining or retrieving information
e. Other changes in physical structure, facilities and equipment
f. Changes in scope and/or nature of services
g. Changes in presence and organisation of quality management mechanisms
h. Changes in ownership and/or affiliation status of hospitals and other facilities
i. Other structural changes in organisation excluding staff (specify)

4.3.2. Staff-oriented changes *(as in 2.1.3b)*

a. Revision of professional roles
b. Multidisciplinary teams
c. Case management
d. Other integration of services
e. Skill mix changes
f. Outcomes improving provider satisfaction
g. Other, specify

4.3.3. Patient-orientated changes *(as in 2.1.3c)*

a. Outcomes facilitating individual patient participation
b. Outcomes facilitating patient group participation
c. Other, specify

4.4. Economic processes and outcomes

In a guidelines implementation project the small extra effort called for to collect prospective data on the costs and benefits of implementing the guideline may be very useful. Reconstructing these data after the project is up and running or has been completed is often complex and unsatisfactory. Therefore we suggest that guideline implementers consider both costs and the type of economic analysis they may want to produce at the planning stage of their project. The outcomes for these analyses usually are identical with main project outcomes. The relevant direct health care costs and benefits most often appear already in parts 4.2 and 4.3; this part of the framework reminds us how important it is to translate these into costs and benefits. Types of economic analysis are presented in part 6.1.2 in detail. For a detailed discussion of economic evaluations, see Chapter 5.

4.4.1. Costs of guideline development

a. Direct costs (expert time, group meeting costs, library costs etc.)
b. Indirect costs

4.4.2. Costs of the intervention

a. Direct cost of disseminating guideline (printing, mailing etc.)

b. Indirect cost of disseminating guideline (upgrading computer systems etc.)
c. Direct cost of implementing guideline (patient leaflets, new drugs etc.)
d. Indirect cost of implementing guideline (lost consultation fees etc.)

4.4.3. Relevant direct health care costs and benefits (e.g. drugs, operations, hospital stays etc.)

4.4.4. Relevant indirect health care costs and benefits (e.g. working days saved etc.)

5. Barriers and facilitators

The same factors may act as either barriers or facilitators for guideline implementation. For example, knowledge about the existence of a guideline helps in using it, while lack of this knowledge completely prevents application. Attitudes toward guidelines may be negative or positive in general; an example of attitude measurement is given in Chapter 8. A more extensive discussion of barriers and facilitators in guideline implementation research is presented in Chapter 6, and a review of studies having looked at various perceived promotive and preventive factors in the implementation of changes in health care is presented in Chapter 9. A library of these studies is available at the CPP website at: www.dsi.dk/projects/cpp.

5.1. Professionals

5.1.1. Knowledge

a. Knowledge about existence of guideline
b. Knowledge about own practice differing from guideline
c. Knowledge for complying with guideline
d. Other, specify

5.1.2. Skills

a. Skills for locating or fetching guideline (e.g. from Internet)
b. Individual skills for complying with guideline
c. Team skills for complying with guideline
d. Organisational competence for complying with the guideline
e. Other, specify

5.1.3. Attitudes

a. Attitudes toward guidelines in general
b. Attitudes toward the guideline in question
    i) Content of guideline
        1. Evidence base or reliability of the guideline
        2. Usefulness of guideline
        3. Other, specify
    ii) Producer of guideline
    iii) Promoter of guideline
    iv) Other, specify

5.2. Patients

5.2.1. Knowledge

a. Knowledge about existence of guideline
b. Understanding of guideline content
c. Other, specify

5.2.2. Skills

a. Understanding recommendations by the professional
b. Ability to follow recommendations by the professional
c. Other, specify

5.2.3. Attitudes

a. Patient's values differ from those of the professional or those in the guideline
b. Relative's values differ from those of the professional or those in the guideline
c. Social group's values differ from those of the professional or those in the guideline
d. Other, specify

5.2.4. Other resources

a. Money
b. Assistance
c. Other, specify

5.3. Environment

5.3.1. Social factors

a. Support for or discouragement of change by others
   i) Colleagues at practice site
   ii) Other members of professional team
   iii) Managers
   iv) Other local health care providers
   v) Opinion leaders
   vi) Patients
   vii) Professional organisations
   viii) Patient organisations
   ix) Others, specify
b. Other, specify

5.3.2. Organisational factors

a. Availability of guidelines at workplace
b. Practicality within existing practice setting or routines
c. Local infrastructures or rules
d. Other, specify

5.3.3. Economic factors

a. Availability/lack of resources (time, personnel etc.)
b. Change in income for provider
c. Changed cost for patient
d. Changed cost for practice organisation
e. Changed cost for health care system
f. Other, specify

6. Methods

6.1. Study design

Guideline implementation studies are resource-intensive, technically rather demanding and require multiprofessional expertise. For these reasons it is advisable to use the strongest possible study design: preferably a randomised controlled trial, a controlled design or a time series analysis with a sufficient number of data points. It is useful to plan in the beginning of the study to combine economic evaluation and/or qualitative approaches to the basic design. More detailed discussions of

methodology in guidelines research are to be found in Chapters 3 (design and statistical issues), 4 (economic evaluation) and 5 (qualitative methods).

6.1.1. Main study design

a. Randomised controlled trials (RCT): any form of randomisation process. This includes alternate allocation and other quasi-random processes.
b. Controlled before and after studies (CBA): e.g. involvement of intervention and control groups other than by random process, and inclusion of baseline period of assessment of main outcomes.
c. Interrupted time series (ITS): study observing a change in trend that is attributable to the intervention.
d. Other, specify

6.1.2. Economic evaluation

a. Cost minimisation analysis: Costs in monetary units, effects assumed to be similar for various types or scales of interventions.
b. Cost-effectiveness analysis: Costs in monetary units, consequences in any measurement of effect of interest. Result in cost per unit of effect.
c. Cost-utility analysis: Costs in monetary units, consequences in any measurement of utility.
d. Cost-benefit analyses: Both costs and consequences measured in monetary units.

6.1.3. Qualitative part(s) of study design

a. Before quantitative study: Questionnaire development, optimising intervention
b. During quantitative study: Studying processes, providing context
c. After quantitative study: Explaining unexpected findings

6.1.4. Other, specify

6.2. Unit of analysis

This item specifies the primary unit of study. The method of analysis may depend on the primary unit: patients/providers allows for classical analysis, practices/hospitals calls for a cluster analysis and episodes of care are analysed in their own manner. See also Chapter 3.

6.2.1. Episodes of care

6.2.2. Patients

6.2.3. Providers

6.2.4. Practices

6.2.5. Hospitals

6.2.6. Communities or regions

6.2.7. Other (specify)

6.3. The number included in the study

The numbers of those included in the study must be stated at least for the main unit of analysis (6.2). When there are several levels of participants (such as professionals within several hospitals), the numbers for each level should be stated. This allows for power calculations for various designs. Numbers need to be given separately for the sample size, the actual number entered and number completing the study according to protocol.

6.3.1. Episodes of care

6.3.2. Patients

6.3.3. Providers

6.3.4. Practices

6.3.5. Hospitals

6.3.6. Communities or regions

6.3.7. Other (specify)

References

Cochrane Effective Practice and Organisation of Care Review Group (EPOC): The Data Collection Checklist. Available at:
http://www.abdn.ac.uk/public_health/hsru/ccepp/index.htm

Eddy DM. Practice policies – what are they? *JAMA* 1990; 263:877-80.

Grimshaw J, Eccles M, Russell IT: Developing clinically valid practice guidelines. *Journal of Evaluation of Clinical Practice,* 1995; 1: 37-48.

Grimshaw J, Russell I. Achieving health gain through clinical guidelines. I: Developing scientifically valid guidelines. *Quality in Health Care* 1993; 2: 243-48.

Institute of Medicine: *Guidelines for clinical practice: From development to use.* Washington DC: National Academic Press, 1992.

Lohr KN: The quality of practice guidelines and the quality of health care. Page 42-52 in: *Guidelines in health care. Report of a WHO-Conference.* Band 104, Schriftenreihe des Bundesministeriums für Gesundheit. Baden-Baden: Nomos Verlag, 1998:

Acknowledgements

The core members of the CPP taxonomy group were Jeremy Grimshaw, Richard Grol, Jim Kahan, Finn Børlum Kristensen, Marjukka Mäkelä, Andy Oxman, Thorkil Thorsen and Michel Wensing. In addition to discussions among this group, insights and suggestions from many other members of the CPP partner projects during the CPP workshops and steering group meetings have been invaluable in developing this framework.

Section B

Methodological issues

Chapter 3

Design and statistical issues in implementation research

Marion K. Campbell, Nick Steen, Jeremy M. Grimshaw, Martin Eccles, Jill Mollison, Carl Lombard

Aims of the chapter

In this chapter we will discuss design and statistical issues which are specific to guideline implementation research. We assume that readers have a basic understanding of research designs and methods. We have not attempted to cover comprehensively general aspects of study design and refer readers to other general texts (for example, Pocock (1983) for discussions of randomised trials, and Cook and Campbell (1979) for discussion of quasi-experimental studies). In this chapter we consider whether evaluations of guideline dissemination and implementation strategies should be pragmatic or explanatory; specific issues concerning what data should be collected, the strengths and weaknesses of different potential designs and statistical issues for sample size calculation and analysis of cluster randomised trials. In addition, the CPP statistics group has produced a number of resources to support sample size calculation of cluster randomised trials.

Introduction

If policy makers are to make evidence-based decisions about guideline implementation, they need information on the effectiveness and cost-effectiveness of different interventions (in different settings for different

targeted clinicians and behaviours), the likely effect modifiers and the resources needed to deliver interventions. In order to provide such information, researchers need to use rigorous designs and methods to allow reliable and valid estimates of the likely effects of alternative interventions to be made. We can then have confidence that the observed effects are both attributable to the interventions studied and generalisable to other contexts. However many existing studies use weak designs or are methodologically flawed with potentially major threats to validity, thereby limiting their value to inform decision making (Bero 1998). The solution to these problems is to conduct rigorous evaluations of guideline implementation strategies. However, such studies are both complex and methodologically challenging.

Explanatory or pragmatic studies of implementation strategies

Schwartz and Lellouch (1967) clarified the distinction between explanatory and pragmatic studies. *Explanatory studies* aim to test whether an intervention is efficacious, that is whether the intervention is effective under ideal conditions. Contextual factors and other effect modifiers are equalised between study groups. Typically they are conducted in highly selected groups of subjects under highly controlled circumstances.

Patients withdrawing from such a study may be excluded from analysis. The narrow inclusion criteria and rigid conduct of explanatory studies limit the generalisability of the results to other subjects and contexts. In contrast, *pragmatic studies* aim to test whether an intervention is likely to be effective in routine practice by comparing the new procedure against the current regimen; as such they are the most useful trial design for developing policy recommendations. The aim is to optimise contextual factors and other effect modifiers in the intervention and study groups, thus approximating normal conditions. In pragmatic studies, the contextual and effect modifying factors therefore become part of the interventions. Such studies are usually conducted on a predefined study population and withdrawals are included within an 'intention to treat' analysis; all subjects initially allocated to the intervention group would be analysed as intervention subjects irrespective of whether they received the intervention or not. Guideline implementation studies should be pragmatic: we are interested in whether guideline dissemination and implementation strategies are effective and efficient in real world settings.

What to measure

Process or outcome measurement?

When planning studies, researchers have to decide what to measure. Classically, researchers have been encouraged to measure both process and outcomes of care. To statistically detect a clinically significant change in outcome of care generally requires greater sample sizes and additional research resources. When evaluating evidence-based guidelines where the link between the process and outcome of care is well understood, it may be possible to measure process of care only.

Baseline measurement of performance

Baseline measures are useful because they provide an estimate of the magnitude of a problem. Low performance scores prior to the intervention may indicate that performance is poor and there is much room for improvement. On the other hand, high performance scores may indicate that there is little room for improvement (ceiling effect). Ideally such measurement should take place in the planning or pilot stage of an implementation trial. In addition, in a randomised or quasi-randomised study, comparing baseline data in the experimental and control groups can provide some reassurance about the adequacy of the allocation process.

Learning and decay effects

Interventions to change professional behaviour may have a gradual rather than an instant effect (learning effect) which may then decline over time (decay effect). It is important to measure change over time to identify whether learning and decay effects are present and to quantify their influence. Guideline researchers should attempt to measure the process or outcome of care serially over time to achieve this. The methodology in this area is still developing, however plotting performance over time will provide a simple and useful way of looking for such effects.

How to measure

Self-reported or objective measurement of performance?

Commonly, researchers have measured performance by asking the professional involved what they have done or intend to do in a specific set of circumstances (for example, at what level of hypertension would you initiate pharmacological treatment?). However there is empirical evidence that self-reports of activity tend to overestimate actual performance (Eccles 1999; Adams 1999). It could be argued that in experimental studies of guideline implementation strategies, this would not be im-

portant if the professionals in the experimental and control groups inflated their estimate by the same degree. However, there is a danger that the intervention may sensitise the professionals in the experimental group about desired practice, potentially leading to an imbalance in the degree to which the experimental and control groups report their behaviour. This could lead to an overestimate of the effect of the intervention: given these concerns, guideline implementation researchers should measure actual performance and not rely on self-report.

Minimal intrusiveness of data collection

Researchers also need to be aware that the methods of data collection may sensitise professionals about desired practice. For example, de Dombal *et al*. (1991) observed that the introduction of structured collection of clinical data by professionals improved performance. If the data collection methods are intrusive, they may lead to improved performance in both the experimental and control groups, potentially leading to an underestimate of the effect of an intervention. Guideline implementation researchers should therefore attempt to use minimally intrusive data collection methods. For example, in the TEMPEST study (see Appendix page 212) patients at risk of deep vein thrombosis were identified from routine data collection systems and evidence of prophylaxis was assessed from the clinical records; the professionals in the study were not sensitised by the ongoing data collection exercise.

Study design

There are a variety of study designs that could be used to evaluate guideline implementation strategies. These vary in the degree to which they allow observed effects to be attributed to the intervention with confidence. In this section, we consider the strengths and weaknesses of different designs.

Observational studies

Observational (or descriptive) studies of single groups may usefully provide greater understanding of the process of behavioural change and generate hypotheses for further testing in rigorous evaluations (Grilli and Lomas 1994). However, they are rarely useful for evaluation because the populations to be compared may differ in characteristics that affect the outcomes being measured – characteristics other than the interventions to be compared. If the evaluator cannot identify or measure these differences, nothing can be done to ameliorate the resulting bias. Even when it is possible to adjust for recognised differences, it is never possible to rule out unrecognised bias with confidence.

Quasi experimental designs

Uncontrolled before and after studies

Uncontrolled before and after studies are weak evaluative designs (Russell 1992) as secular trends or sudden changes make it difficult to attribute observed changes to the intervention (Cook and Campbell 1979). Furthermore, there is some evidence to suggest that the results of uncontrolled before and after studies may overestimate the effects of interventions (Lipsey 1993; Soumerai 1989).

Controlled before and after studies

Controlled before and after studies incorporate a non-randomised control group that will experience trends and changes similar to those of the study population, thus overcoming some of the difficulties of uncontrolled before and after studies. Control professionals need to be chosen carefully to ensure comparability with study professionals and should have similar baseline characteristics and performance (Effective Health Care 1994). However the usefulness of controlled before and after studies is limited because the estimate of effect cannot be attributed to the intervention with confidence due to the non-randomised control group. In many circumstances, where a controlled before and after design is proposed, a randomised trial could be as easily undertaken and provide a more reliable estimate of effect.

Interrupted time series

Time series analyses are useful in guideline implementation research for evaluating the effects of interventions when it is difficult to randomise or identify an appropriate control group (for example, following the dissemination of national guidelines or mass media campaigns). Time series analysis can be used to detect whether an intervention has had an effect significantly greater than the underlying trend (Cook and Campbell 1979). This increases the confidence with which the estimate of effect can be attributed to the intervention although the design does not provide protection against the effects of other events occurring at the same time as the study intervention which might also improve performance.

Randomised trials

Patient randomised trials

Randomised trials are rightly considered the most robust method of assessing health care innovations (Cochrane 1972). Randomised trials estimate the impact of an intervention through direct comparison with a randomly allocated control group that either receives no intervention or an alternative intervention. The randomisation process ensures that,

all else being equal, both known and unknown biases are distributed evenly between the trial groups.

When evaluating behavioural change strategies, however, simple (patient) randomised trials may be less robust. There is a danger that the treatment offered to control patients will be contaminated by doctors' experiences of applying the intervention to patients receiving the experimental management, with the result that the evaluation may underestimate the true effects of strategies. For example, Morgan and colleagues (1978) undertook a study of computerised reminders for antenatal care. They chose to randomise patients between a control group and an experimental group for whom any non-compliance by the doctor generated an automatic reminder from the computer-based medical record system. Compliance in experimental patients rose from 83% to 98% within six months, while compliance in control patients rose from 83% to 94% in 12 months. The results suggest that the intervention had a significant (if delayed) effect on the management of control patients.

Cluster randomised trials

To overcome these problems, it is possible to randomise groups of professionals. In such circumstances we would still want to collect data about the process and outcome of care at the individual patient level. Such trials, which randomise at one level and collect data from a different level, are known as cluster randomised trials.

Whilst cluster randomisation to a large extent overcomes the problem of contamination in patient randomised trials, it has implications for the planning, conduct and analysis of studies. A fundamental assumption of the patient-randomised trial is that the outcome for an individual patient is completely unrelated to that for any other patient – they are said to be 'independent'. This assumption is violated, however, when cluster randomisation is adopted, because patients within any one cluster are more likely to respond in a similar manner. For example, the management of patients in a single hospital is more likely to be consistent than management across a number of hospitals. The primary consequence of adopting a cluster randomised design is that it is not as statistically efficient as a patient randomised design; it has lower statistical power than a patient-randomised trial of equivalent size (Donner 1998).

Despite the added complexity, cluster randomised trials provide the optimal design for guideline implementation studies. In the next section, we discuss the advantages and disadvantages of different approaches to the conduct of cluster randomised trials.

Possible types of cluster randomised trials

Two-arm trials

In two-arm cluster trials, groups of professionals are randomised to study or control group. Such trials are relatively simple to design and operationalise, and they maximise power (half the sample is allocated to the intervention and half to the control). If two arm trials are used to evaluate a single intervention against control, however, they only provide limited information about the effectiveness of the intervention within a single setting. They do not provide information about the relative effectiveness of different interventions within the same setting. Two arm trials can also be used to compare two different interventions, but do not provide information about the effect of either intervention against a control. For these reasons, implementation researchers should consider other design options which will allow them to compare interventions head-to-head and against a control. The other design options below are in general more complex to use.

Multiple arm trials

The simplest extension to the two-arm trial is to randomise groups of professionals to more than two groups (for example, two or more study groups and a control group). Such studies are relatively simple to design and use, and allow head-to-head comparisons of interventions or levels of intervention under similar circumstances. These benefits are, however, compromised by a loss of statistical power; for example, to achieve the same power as a two-arm trial, the sample size for a three-arm trial needs to be increased by up to 50%.

Factorial designs

Factorial designs allow the comparison of more than one intervention with reduced loss of power compared with multiple arm trials. For example in a 2×2 factorial design evaluating two interventions against control, participants are randomised to each intervention (A and B) independently (see Table 3.1). In the first randomisation, the study participants are randomised to intervention A or control. In the second randomisation, the same participants are randomised to intervention B or control. This results in four groups: no intervention, intervention A only, intervention B only, both intervention A and B. During the analysis of factorial designs, it is possible to undertake independent analyses to estimate the effect of the interventions separately (Cochran and Cox 1957); essentially this design allows the conduct of two randomised trials for the same sample size as a two arm trial. However these trials are more difficult to operationalise and analyse, they provide only limi-

Table 3.1. Diagrammatic representation of a factorial design

Intervention B	*Intervention A* *No*	*Yes*
No	Group 1 receive neither intervention	Group 2 receive intervention A only
Yes	Group 3 receive intervention B only	Group 4 receive both interventions

ted power for a direct head-to-head comparison of the two interventions and the power is diminished if there is interaction between the two interventions.

Balanced incomplete block designs

In guideline implementation research, there are also a number of non-specific effects which may influence the estimate of the effect of an intervention. Currently, these non-specific effects are lumped together and termed the 'Hawthorne effect'. If these are imbalanced across study groups in guideline implementation trials, the resulting estimates of effects may be biased.

Balanced incomplete block designs can be used to equalise such non-specific effects and thereby minimise their impact (Cochran 1957). For example, the COGENT study (see page 214) used a 2×2 balanced incomplete block design. Study doctors were randomly allocated between two groups. One group received computerised guidelines for the management of asthma and provided control data for the management of angina. The other group received the computerised guidelines for the management of angina and provided control data for the management of asthma (see Table 3.2). As doctors in both groups are subject to the same level of intervention, the Hawthorne effect should be equalised across the two groups.

Table 3.2. Example of balanced incomplete block design

	Asthma	*Angina*
Doctor group 1	Intervention	Control
Doctor group 2	Control	Intervention

Statistical issues in cluster randomised trials

We have argued that cluster randomised trials are the 'gold standard' for guideline implementation studies. However there are a number of statistical issues particular to the planning, conduct and analysis of these trials, which we outline below.

The intracluster correlation coefficient

Within cluster randomised trials, patients within any one cluster are often more likely to respond in a similar manner. A statistical measure of this intracluster dependence is known as the 'intracluster correlation coefficient' (ICC) which is based on the relationship of the between to within-cluster variance (Donner and Koval 1980). For example, in a study which randomised by hospital, the ICC would be high if the management of patients within hospitals was very consistent but there was wide variation in practice across hospitals.

Estimates of the likely size of ICCs are essential for both sample size calculation and analysis of cluster trials. There is little empirical evidence available, however, on their likely size, and on what factors influence their magnitude. As part of the Concerted Action, empirical estimates of ICCs were calculated from a number of implementation datasets. A selection of these is presented in Table 3.3, and a full listing can be freely accessed and downloaded from the Health Services Research Unit's website: http://www.abdn.ac.uk/public_health/hsru/icc.html.

Whilst their numerical values may appear small, even small ICCs can have a significant effect on sample size requirements (see Table 3.3, page 66).

In primary care settings, the ICCs for process variables appear to be of an order of magnitude higher than those of outcome variables; estimates for process variables from primary care were of the order of 0.05 to 0.15, whereas ICCs for outcome variables were generally lower than 0.05. This observation may be explained by the greater biological variability intrinsic in measures of patient outcome compared with measures of physician behaviour. For example, although physicians may be consistent in their use of a treatment, patients will vary in their compliance with and response to that treatment. ICCs for process variables from the secondary care setting were of an order higher than those taken from primary care (around 0.3 in secondary care as compared with 0.05 to 0.15 in primary care).

On the basis of the current estimates, it would be reasonable for a researcher planning an implementation study within primary care in the United Kingdom to assume an ICC of the order of 0.1 for process variables, and an ICC of less than 0.05 for outcome variables. Further estimates of ICCs are required for the planning of secondary care studies and studies outside the United Kingdom. As these become available they will be added to the database of ICCs.

Table 3.3. Empirical estimates of intracluster correlation coefficients

Unit of allocation	*Outcome*	*No. of clusters*	*Average cluster size*	*ICC*	*Data set*
Process measures:					
Practice	Recording of previous medical history	62	31	0.06	North of England Study (1992)
Practice	Child examined	62	31	0.01	North of England Study (1992)
Practice	Recording of investigations	62	31	0.04	North of England Study (1992)
Practice	Recording of advice given	62	31	0.07	North of England Study (1992)
Practice	Appropriateness of referral score	63	5	0.04	Aberdeen GRIP study
Practice	No. of annual referrals	68	4	0.24	Aberdeen GRIP study
Practice	Appropriateness score for content of referral letter	63	5	0.15	Aberdeen GRIP study
Practice	No. of annual referrals	58	15	0.12	Aberdeen URGE study
Practice	No. of visits to practice prior to referral	58	15	0.11	Aberdeen URGE study
Obstetric unit	Use of antibiotics in caesarian section deliveries(within 6 hours of surgery)	25	28	0.30	Wyatt et al (1998)
Obstetric unit	Use of ventouse in instrumental delivery	25	30	0.21	Wyatt et al (1998)
Obstetric unit	Use of prophylactic corticosteroids in pre-term deliveries	25	6	0.22	Wyatt et al (1998)
Obstetric unit	Use of polygylcolic acid sutures for deep repair after perineal tears or episiotomies	25	28	0.66	Wyatt et al (1998)
Hospital	Use of out-patient endometrial biopsy as method of endometrial sampling	12	100	0.24	Gynaecology Audit Project Scotland
Hospital	Use of semen analyses (≥2) in management of infertility	12	125	0.18	Gynaecology Audit Project Scotland
Outcome measures:					
Practice	Patients with controlled hypertension	18	49	0.064	Fahey & Peters (1996)
Practice	SF36 – physical functioning score	64	12	0.05	Aberdeen URGE study
Practice	SF36 – role physical score	64	12	0.01	Aberdeen URGE study
Practice	SF36 – role emotional score	64	12	0.008	Aberdeen URGE study
Practice	SF36 – social functioning score	64	12	0.02	Aberdeen URGE study
Practice	SF36 – mental health score	64	12	0.007	Aberdeen URGE study
Practice	SF36 – energy and vitality score	64	12	0.03	Aberdeen URGE study
Practice	SF36 – pain score	64	12	0.03	Aberdeen URGE study

Sample size calculation

Standard sample size formulae also assume that the outcome for each patient is independent: if these formulae are used for cluster randomised trials, they will result in under-powered studies. To accommodate for this, sample size estimates should be inflated to adjust for the clustering effect. For completely randomised designs, standard sample size estimates should be inflated by a factor of:

$$1+(\bar{n}\text{-}1)\,\rho$$

where $\bar{n}$ is the average cluster size, and ρ is the estimated ICC (Donner *et al.* 1981). This inflation factor is also known as the 'design effect'. As both the ICC and the cluster size influence the inflation required, the design effect can be considerable even when the ICC is small (if the average cluster size is large).

The extra numbers of patients required can be achieved by increasing either the number of clusters in the study or the number of patients per cluster; however, increasing the number of clusters is the most efficient method (Diwan 1992). This is shown in Table 3.4., where, for example, the increase from 50 to 100 patients per practice (assuming an ICC of 0.1) only reduces the number of practices required by two. The number of clusters recruited to a trial also influences the types of analysis which may later be undertaken. Some techniques cannot be used if too few clusters are available, for example matched analysis becomes problematic with fewer than 6 pairs of clusters (Hayes *et al.* 1997).

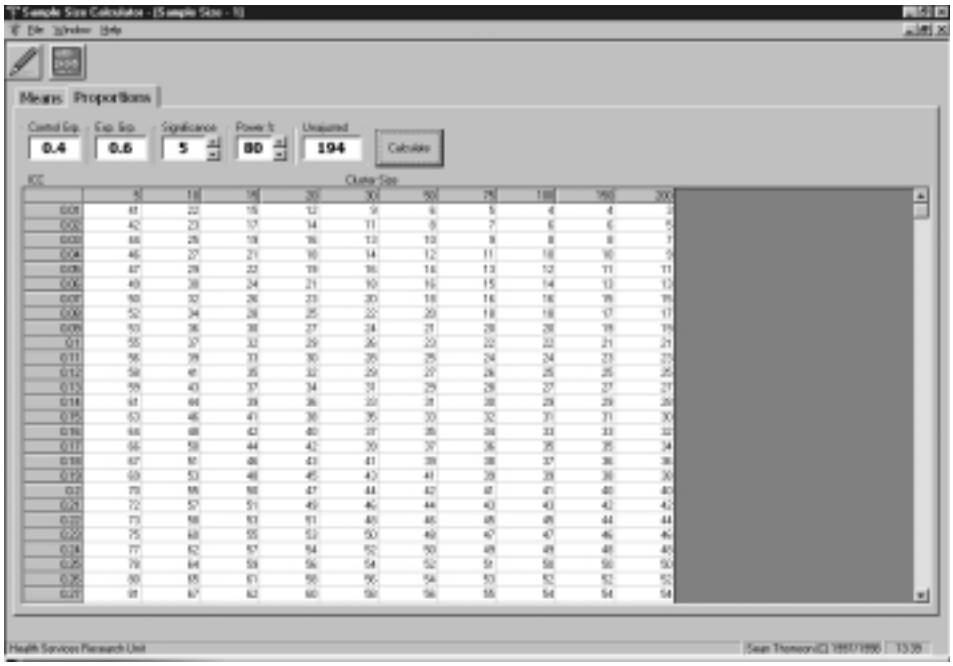
Table 3.4. Example of impact of randomisation by cluster on sample size

Example: What is the sample size required to detect a difference of 20% in compliance with a guideline between an intervention and control group (from 40% to 60%, with 80% power and a 5% significance level)? Assuming no clustering a total of 194 patients would be required.

	No. of patients per cluster			
	10	***20***	***50***	***100***
ICC	*Number of clusters (total number of patients)*	*Number of clusters (total number of patients)*	*Number of clusters (total number of patients)*	*Number of clusters (total number of patients)*
0.01	22 (220)	12 (240)	6 (300)	4 (400)
0.05	30 (300)	20 (400)	14 (700)	12 (1200)
0.10	38 (380)	30 (600)	24 (1200)	22 (2200)
0.15	46 (460)	38 (760)	34 (1700)	32 (3200)

A sample size calculator has been developed for the planning of cluster randomised trials as part of the CPP Concerted Action. It allows

Figure 3.1. Sample size calculator



the researcher to calculate sample size requirements for varying estimates of ICC and varying cluster sizes (Figure 3.1). It can be accessed and downloaded from the Health Services Research Unit's website: http://www.abdn.ac.uk/public_health/hsru/sampsize.htm

Strategies for improving power

Logistic or resource constraints often limit the number of clusters available for randomisation in guideline implementation trials. As a result, many implementation trials have limited statistical power. There are a number of design features, however, that could enhance power, which we discuss below.

Level of randomisation

Researchers need to think carefully about the level at which to randomise, as there is a trade off between the risk of contamination and number of clusters available. For example, in the TEMPEST study (see Appendix A), randomisation at the level of the hospital would have minimised the risk of contamination but would have dramatically increased the number of hospitals required. This would have substantial logistical implications, limiting the feasibility of the study. In contrast, randomi-

sation at the level of the hospital ward would decrease the number of hospitals required but increase the risk of contamination because of shared policies and cross cover of staff within the same hospital. The researchers had to trade off considerations of contamination and increased logistical problems associated with recruiting more hospitals. They chose to randomise at the level of clinical service (i.e. all general surgical wards within one hospital were randomised as one cluster). Whilst there was possible risk of contamination between clinical services, this was considered to be of an acceptable level.

Stratification/Matching

Randomisation of small numbers of clusters may lead to chance imbalances between the arms of a trial on factors such as baseline performance or characteristics of the clusters (for example, type of hospital) which may influence the observed intervention effect. To reduce the chance of this happening, it is possible to pair-match or stratify the clusters on some of the important variables, for example hospital size (Donner 1998). If there is high correlation between the outcomes of a matched pair or stratified unit, this will increase the power of the study. The analysis of these study designs is more complex. Alternatively, a less powerful analysis ignoring the matching or stratification could be undertaken.

Frequency of measurement

Another method that may help increase power is to take repeated measurements on subjects, or repeated measures on independent cohorts of subjects, for example collecting data before and after the intervention. For example, Duffy *et al* (1992) showed that the incorporation of antecedent data considerably increased the power of large public health trials.

Analysis of cluster randomised trials

Within this section, we discuss analysis of cluster randomised trials. As no CPP partner projects were completed, we have used the Urology Referral Guidelines Evaluation (URGE) study to illustrate different approaches to analysis (see Box 3.1, page 70).

Within this section, we focus on the evaluation of the effectiveness of the intervention for benign prostatic hyperplasia (BPH) patients only. Data for a single outcome will be used – waiting time from the date of patient referral to first appointment at hospital. Waiting time was measured in days and was found to have a skewed distribution that was log transformed to normality. Therefore geometric means are quoted

Box 3.1.
The Urology Referral Guidelines Evaluation (URGE) study

The URGE study aimed to evaluate the effectiveness of a guideline-based open access 'fast track' investigation service for two common urological problems – benign prostatic hyperplasia (BPH) and microscopic haematuria. General practices were randomly allocated to two groups; one group received guidelines for the appropriate referral of BPH patients for the open access 'fast track' system whilst the other group acted as a control for BPH patients (but did receive guidelines for another condition, microscopic haematuria).

Data were collected on two cohorts of patients, one referred before and another referred after the introduction of the fast track service (an indicator of baseline performance). Data were collected on pre referral general practice management, hospital and general practice care following referral and patient outcome.

throughout; the effect sizes and the corresponding 95% confidence intervals relate to the ratio of mean waiting time in the intervention group compared to the control group. Data were available on 513 patients (211 before and 312 after the introduction of the fast-track service) referred from 54 general practices from the Grampian region of Scotland.

There are two main approaches to the analysis of cluster randomised trials: analysis at the cluster level or analysis at the patient level.

Traditionally, analysis has been focused at the cluster level; however, recent advances in statistics have led to the development of techniques which can incorporate the patient level data. Within each approach, simple analyses such as t-tests or more complex approaches such as regression analyses may be undertaken. Both allow the effect of the intervention to be tested; however, only complex analyses allow adjustment for potential covariates, such as baseline performance.

Analytic methods for each approach are described below. It should be noted that these methods are appropriate for completely randomised designs. Readers should refer to more detailed texts, for example Murray (1998) for discussion of the appropriate methods to analyse stratified or matched designs.

Cluster level analysis

The traditional approach to the analysis of cluster randomised trials has been to calculate a summary measure for each cluster, such as a cluster

mean or proportion. Because each cluster then provides only one data point, the data can be considered to be independent, allowing standard statistical tests to be used.

For example, within the URGE trial, the mean waiting times post intervention for each general practice could be calculated (when different patients are included pre and post, only post data comparisons can be made using simple analyses) (see Table 3.5). The overall group means can then be compared using a standard t-test resulting in a significance of $t_{48} = 3.99$, $p = 0.0001$. This results in an effect size of 0.65 (95% CI: 0.53 to 0.81), in other words the waiting time was on average 35% less in the guideline group. When the size of the clusters vary widely it is preferable to carry out a weighted t-test, using cluster sizes as the weights (Kerry 1998). This weighted analysis returns an effect size of 0.65 (95% CI: 0.54 to 0.78), with a significance of $t_{48}=4.72$, $p = 0.00001$.

Table 3.5. Post intervention mean waiting times (days) per practice

Practice	*Intervention*	*Control*
Practice A	43.6	.
Practice B	61.6	.
Practice C	.	83.9
Practice D	.	68.7
. . .		
Overall mean	***37.8***	***57.9***

Standard statistical techniques such as multiple regression can also be used when data have been summarised at a cluster level. These analyses, however, can only adjust for cluster-level covariates.

Whilst these cluster level approaches overcome the problem of the non-independence of the data, they are in general not statistically efficient. They are efficient in the particular case of the analysis of continuous outcomes when there is no variation in cluster size (Donner 1998).

Patient level analysis

Recent developments in the statistical field now allow all the patient level data to be utilised, whilst accounting for the intracluster correlation; thus increasing the statistical power of the analysis.

Adjustments can now be made to simple statistical tests to account for the clustering effect. For example, test statistics based on Chi-squared or F-tests should be divided by the design effect (as described in the section on sample size), while test statistics based on the t-test or the z-test should be divided by the square root of the design effect (Murray 1998). Adjustments for these and other tests such as nonparametric tests are discussed by Donner and Klar (1994).

In the URGE study, the mean waiting time post-intervention in the guideline group was 39.4 days and 60.6 days in the control group. If the clustering effect had been ignored and a standard t-test performed, the analysis would have resulted in a t-value for the difference between groups of 5.11 (with a highly significant p-value of 0.0000001 based on 310 degrees of freedom), and the resulting effect size would have been 0.65 (95% CI, 0.55 to 0.77).

The design effect for the time to first appointment outcome within the URGE trial was 1.56; hence the revised t-value adjusting for clustering is calculated:

$$\frac{\text{t-value}}{\sqrt{(\text{design effect})}} = \frac{5.11}{\sqrt{(1.56)}} = 4.09$$

resulting in a revised significance level of 0.00003. The 95% confidence interval can also be adjusted for clustering. The revised 95% confidence interval is 0.52 to 0.80.

Despite a highly significant difference in waiting times between the groups, this example illustrates the impact of clustering on the significance of trial results. If clustering had been ignored, the analysis would have returned a spuriously low p-value and overly narrow confidence intervals, over-emphasising the impact of the intervention.

Similarly, there have been advances in the development and use of new modelling techniques to incorporate patient level data such as mixed linear models, hierarchical linear modelling and generalised estimating equations. These modelling techniques allow the inherent correlation within clusters to be modelled explicitly, and thus a 'correct' model can be obtained.

The aim of statistical modelling is to identify the main factors that explain variation in the outcome. In the URGE study factors other than the intervention might also explain variation in the waiting time, for example patient and practice characteristics. When analysing guideline implementation trials, the primary aim of modelling is to adjust for the effect of such covariates before the effect of the intervention is tested rather than to maximise the proportion of variation explained.

An analysis plan or analysis strategy should be developed before any analysis is undertaken to ensure that the modelling is hypothesis-led rather than data-driven. The *a priori* model-fitting analysis strategy should identify:

- the covariates which are to be considered for inclusion in any modelling approach to analysis

- the order in which confounding variables are to be considered for inclusion to the model with the intervention variable fitted last (or an 'intervention × phase' interaction if pre- and post-measurements have been taken. (Cook and Campbell 1979).

An example of a model-fitting analysis strategy which could have been used for the URGE data is displayed in Box 3.2.

Box 3.2.
Example of model-fitting strategy
- phase (pre/post) – design variable
- practice size – covariate
- intervention (guideline/no guideline)
- intervention x phase interaction

Multilevel modelling was undertaken for the URGE study using the software package *MLWin,* developed by the Institute of Education in London. As outlined above, an *a priori* model-fitting analysis strategy was developed which identified the order in which covariates were to be included in the model. Only after all covariates were included in the model was the effect of the 'intervention x phase' interaction examined. After adjustment for the pre-identified covariates, the interaction remained significant. The effect size estimated from the multilevel model was 0.70 (95% CI: 0.55 to 0.91). The resulting t-ratio was t=2.71, df=307, P=0.004. This indicates that, when all the data are used in the analysis, the waiting time was on average 30% less in the guideline group compared with the control group

An in-depth discussion of all the available modelling methods is beyond the scope of this chapter. Researchers should refer to specific texts such as Murray (1998) for a general introduction to possible methods or Kreft & de Leeuw (1998) for discussion of multilevel models. Similarly, a range of statistical software packages is available for the analysis of clustered data sets. A discussion of the more common packages can be found on the multi-level modelling website: http://www.ioe.ac.uk/multilevel/. For a discussion of generalised estimating equations, readers should refer to Burton *et al.* (1998).

These modelling techniques adjust well for clustering and are most statistically efficient as all the available data is utilised. They also allow adjustment for both cluster-level and patient-level covariates. These types of analyses are more computationally intensive, however, and require greater statistical expertise both in the execution of the proce-

dures and in the interpretation of the results. It is recommended that statistical advice is sought early in the planning of the evaluation.

Reporting

General standards are available for the reporting of randomised controlled trials, through the CONSORT statement (Begg 1996), and should be adopted where appropriate. Currently the CONSORT statement does not include the specific details required for the appropriate reporting of cluster randomised trials. Additional information should be included when reporting cluster randomised trials (Table 3.6)

Table 3.6. Additional reporting information for cluster randomised trials

Section	*Additional information required*
Design	unit of allocation justification for choice of cluster design
Sample size	estimate of ICC average cluster size design effect [a] number of clusters required
Analysis	level at which analysis was undertaken justification for choice of level for analysis analysis technique adopted (including indication of how clustering was accounted for)

a) desirable but not essential, as it can be calculated from the ICC and the cluster size

KEY MESSAGES

- Cluster randomised trials are the optimum evaluative study design for guideline implementation research
- Sample size calculations should take account of the cluster randomised design
- Study power may be increased by the use of matching, stratification or repeated measures within the cluster randomised design
- Analysis should take account of the cluster randomised design
- The CPP programme has produced a number of tools to aid the planning of guideline implementation studies
- Expert statistical advice should be sought when planning guideline implementation studies

References

Adams AS, Soumerai SB, Lomas J, Ross-Degnan D. Evidence of self report bias in assessing adherence to guidelines. *International Journal for Quality in Health Care* (1999, in press)

Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I et al. Improving the quality of randomized controlled trials: the CONSORT statement. *JAMA* 1996; 276: 637-9.

Bero L, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote implementation of research findings by health care professionals. *BMJ* 1998; 317: 465-468.

Burton P, Gurrin L, Sly P. Extending the simple linear regression model to account for correlated responses: an introduction to generalized estimating equations and multi-level mixed modelling. *Statist Med* 1998; 17: 1261-1291.

Cochran WG , Cox GM. *Experimental design.* 2nd ed. New York: Wiley, 1957.

Cochrane AL. *Effectiveness and efficiency: random reflections on health services.* London: Nuffield Provincial Hospitals Trust, 1972

Cook TD, Campbell DT. *Quasi-experimentation: design and analysis issues for field settings.* Chicago: Rand McNally, 1979.

De Dombal FT, Dallos V, McAdam WA. Can computer aided teaching packages improve clinical care in patients with acute abdominal pain. *BMJ* 1991; 302: 1495-7.

Diwan VK, Eriksson B, Sterky G, Tomson G. Randomization by Group in Studying the Effect of Drug Information in Primary Care. *Int J Epidem* 1992; 21: 124-130.

Dobson AJ. *An introduction to statistical modelling.* London: Chapman & Hall, 1983.

Donner A. Some aspects of the design and analysis of cluster randomization trials. *Appl Statist* 1998; 47: 95-113.

Donner A, Birkett N, Buck C. Randomization by Cluster, Sample Size Requirements and Analysis. *Am J Epidem* 1981; 114: 906-915.

Donner A, Klar N. Methods for comparing event rates in intervention studies when the unit of allocation is a cluster. *Am J Epidem* 1994; 140: 279-89.

Donner A, Koval JJ. The estimation of intraclass correlation in the analysis of family data. *Biometrics* 1980; 36: 19-25.

Duffy SW, South MC, Day NE. Cluster randomisation in large public health trials: the importance of antecedent data. *Statist Med* 1992; 11: 307-316.

Eccles M, Ford G, Duggan S, Steen N. Are postal questionnaire surveys of reported activity valid: an exploration using general practitioner management of hypertension in older people. *British Journal of General Practice* 1999; 49: 35-38

Effective Health Care. *Implementing clinical guidelines. Can guidelines be used to improve clinical practice?* Effective Health Care Bulletin No 8. 1994. Leeds: University of Leeds.

Fahey TP, Peters TJ. What constitutes controlled hypertension?Patient based comparison of hypertension guidelines. *BMJ* 1996; 313: 93-6.

Grilli R, Lomas J. Evaluating the message: the relationship between compliance rate and the subject of practice guideline. *Med Care* 1994; 32: 202-13.

Hayes R, Wawer M, Gray R, Whitworth J, Grosskurth H, Mabey D and HIV/STD trials workshop group. Randomised trials of treatment for HIV prevention: report of an international workshop. *Genitourin Med* 1997; 73: 432-443.

Kerry SM, Bland JM. Analsyis of a trial randomised in clusters. *BMJ* 1998; 316: 54.

Kreft I, de Leeuw J. *Introducing multilevel modeling.* London: SAGE, 1998.

Lipsey MW, Wilson DB. The efficacy of psychological, educational and behavioural treatment. Confirmation from meta-analysis. *American Psychologist,* 1993; 48: 1181-1209.

Morgan M, Studney DR, Barnett GO, Winickoff RN. Computerized Concurrent Review of Prenatal Care. *Quality Review Bulletin* 1978; 4: 33-36.

Murray DM. *The design and analysis of group randomised trials.* Oxford: Oxford University Press, 1998.

North of England Study of Standards and Performance in General Practice (). Medical audit in general practice: effects on doctors' clinical behaviour and the health of patients with common childhood conditions. *BMJ*, 1992; 304: 1480-8.

Pocock SJ. *Clinical trials: a practical approach.* New York: Wiley, 1983.

Rice N, Leyland. Multilevel models: applications to health data. *JHSRP* 1996; 1: 154-164.

Russell IT, Grimshaw JM. The effectiveness of referral guidelines: a review of the methods and findings of published evaluations. In: Coulter A and Roland M (eds). *Referrals from General Practice.* Oxford, Oxford University Press, 1992.

Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in clinical trials. *Journal of Chronic Diseases* 1967; 20: 637-648.

Soumerai SB, McLaughlin TJ, Avorn J. Improving drug prescribing in primary care: a critical analysis of the experimental literature. *Milbank Q* 1989; 67: 268-317.

Wyatt JC, Paterson-Brown S, Johanson R, Altman DG, Bradburn MJ, Fisk NM. Randomised trial of educational visits to enhance use of systematic reviews in 25 obstetric units, studying 4508 women. *BMJ*, 1998; 317: 1041-6.

Chapter 4

Economic evaluation of guideline implementation strategies

Emma McIntosh

Aims of the chapter

- Provide an introduction to the rationale for the economic evaluation of implementation strategies
- Examine the specific economic evaluation requirements within CPP
- Outline the general issues involved in costing within economic evaluation and more specifically examine those costs of particular relevance within implementation strategies
- Examine benefit assessment in economic evaluation and more specifically the issues arising within implementation studies
- Summarise the techniques for bringing costs and benefits together within an economic evaluation framework
- Describe the balance sheet approach to the economic evaluation of implementation strategies

Introduction

The overall objective of the "Changing Professional Practice" (CPP) project is the study of how to transfer the results of clinical research into clinical practice. One subgoal is to learn how best to implement clinically evidence-based guidelines. This subgoal will be attained in an inter-active process between empirical findings from national experiments on the one hand and support from scientists who study the theo-

retical implications of implementing clinical guidelines on the other. The result being the development of a multi-disciplinary theoretical framework aimed at future research in the field. This chapter will summarise the health economics component of this multi-disciplinary theoretical framework.

The role for economic evaluation within guideline implementation research

Clinical effectiveness is an important factor when addressing the issue of CPP. However, if clinical effectiveness is seen as an isolated end, and not within the scope of its resource implications, the outcome of the changed clinical practice might indeed be a health gain for the specific patient group involved; but seen as a whole, the population might suffer from a total loss of health. Thus, whilst the inclusion of cost considerations may be seen initially as secondary to the effectiveness of the guidelines themselves, this latter perspective shows that cost-considerations have a direct effect on health and thus should be considered at the early stages of evaluation. Therefore, the consideration of efficiency and cost-effectiveness is crucial when deciding how best to change clinical practice. Rosser *et al.* (1993) surveyed Ontario family doctors about their knowledge of lipid lowering guidelines and while 78% of the doctors indicated that they complied with the guidelines, further questioning revealed that only 5% of the respondents actually followed them. Such a finding has major cost implications. When resources have been invested to develop a guideline and yet there is such low usage or implementation, this can be seen as a waste of valuable resources which could have been put to an alternative, more beneficial use.

It can be seen that there are two contributions economics can make in guideline research. Firstly, one can use economics to consider whether it is worth developing a guideline at all and secondly if it has been decided to develop a guideline, it is important to incorporate considerations of cost into the development, dissemination and implementation of the guideline itself. Addressing the first point involves prior consideration of the effect of the latter point, namely the costs and benefits of implementing the guideline. It is this issue which this chapter is concerned with.

Economics in guideline implementation research to date

There have been few published economic evaluations of guideline implementation studies. However, a number of studies have published equivalent cost and benefit data in their evaluation of implementation strategies. Russel and Suarez-Almazor (1997) in their study on clinical

practice guidelines for radiography of the lumbar spine found that if published guidelines to reduce the utilisation of radiographs had been followed, more rather than fewer x-ray studies would have been carried out. This is an example where the additional costs of implementing guidelines would have actually given rise to additional costs with no benefit (or even dis-benefit) to patients and on grounds of efficiency would be extremely cost-*in*effective. This study concluded that a more restricted and cost-efficient set of guidelines should be proposed. On the other hand, Vollman *et al*. (1998) show how the implementation of clinical guidelines gives rise to significant savings in material costs in the health care system.

Aucott *et al.* (1996) evaluated the effects of an intensive intervention to implement local guidelines for cost-effective management of hypertension on medication use, and cost, blood pressure control, and other resource use. A control group received guidelines and usual education for the cost-effective management of hypertension and the intervention group received guidelines plus intensive guideline-based education and supervision. The study showed that intensive implementation of guideline-based education and supervision was associated with an increased use of guideline medications, decreased use of costly alternative agents, and no reduction in the measured outcomes of care. While the paper states that there was no increased use in other measured resources in the intervention group including the number of outpatient laboratory services, clinic visits or hospitalisations, the paper does not report the actual cost of implementing the guidelines. The worst case scenario would be that the costs involved in implementing the guidelines are higher than the cost savings incurred by using the guidelines. However, it may be the case that the costs of implementing the guidelines were more than offset by the resulting savings. A paper by O'Connor *et al.* (1996) also reported positive findings in a study evaluating a clinical guideline. But again the authors do not report the actual costs of implementation, hence it is difficult to tell whether the costs of the implementation programme could possibly have outweighed the resulting cost savings. Information on the implementation costs of both these guidelines would have provided the reader/policy maker with much more useful data.

A paper by Lauterbach (1998) states that clinical practice guidelines can be used to achieve optimal utilisation of scarce resources for the medical management of defined patient groups. However, Lauterbach states that this does not mean that the total amount of resources allocated to these patient groups will be reduced automatically. The paper discusses examples of different health economic consequences for guideline implementation, especially with regards to the cost-effectiveness

to therapy and to the total health care budget. Lauterbach states: "The development of evidence-based guidelines without accounting for cost-effectiveness of therapy is not reasonable as cost-effectiveness is implied in the definition of guideline objectives. Practice guidelines that are not taking into consideration health economic data may diminish the cost-benefit relation of therapy".

Economic evaluation and CPP

This section outlines the specific issues involved with the CPP collaboration and economic evaluation. Firstly, it should be pointed out, that as has been found with many of the other contributing disciplines to this CPP partnership, there is no widely accepted successful way to incorporate economic considerations into guideline implementation research. In fact, it can be said that the apparent complexity of incorporating costs into guidelines and into deciding whether they are worth developing has clearly been an obstacle to doing so. Whilst there is no agreed method as such, there has however been a consensus on the principles underlying economic evaluation, why they are important and why they should be included in implementation research. In addition to this, an approach, termed the 'balance-sheet' approach (McIntosh, Donaldson & Ryan 1999), to the economic evaluation of guidelines has proven a useful framework for outlining the costs and benefits of the implementation research. This has provided a useful starting point, as the use of such an approach will enable, at the very least, identification of the relevant margins of change. Measurement and valuation of these marginal changes is the more complex issue.

Economic evaluation

Economic evaluation should be seen as a decision making framework which renders the costs and benefits of any intervention or service explicit. In doing so, informed decisions can be made about the allocation of resources to various programmes. The outcome or benefit measure chosen will decide the actual 'type' of economic evaluation carried out, if a number of outcome measures are chosen, a variety of economic evaluations may be performed and the results used as an aid to decision making. By ensuring that the opportunity costs of programmes are minimised this should ensure the maximisation of well-being to society given limited resources.

In the next section, issues arising in the identification and valuation of costs are considered. The important concepts in costing exercises, such as: opportunity cost; the margin, discounting; and sensitivity ana-

lysis were briefly touched on in the CPP sessions and they are outlined more fully here. Attention will then turn to benefit assessment. Two issues are discussed: what is it we are trying to measure; and how can we measure such factors? It is argued that despite the emphasis on health outcomes in recent years, benefit assessment in health economics should also consider non-health outcomes and process attributes. This is especially pertinent within guidelines as often it is the process of care which is most significantly affected or most easily measured as an indication of success of implementation. Attention is then given to methods of valuing health outcomes, non-health outcomes and process attributes. Following this, it is shown how costs and benefits can be brought together within a formal economic evaluation. The special case of guideline implementation however, may not be conducive to this 'combining' of costs and benefits in such a decision-science manner; thus, the technique of the balance sheet approach is proposed and outlined.

An important point to note is that not all possible costs and benefits have to be or indeed should be included within every economic evaluation. It depends upon the perspective of the evaluators as to which are the 'relevant' costs and benefits. A purchaser may only be interested in health service costs and benefits whilst an evaluation from the patients perspective would only be interested in patient costs and benefits not those associated with the health service. A societal perspective would consider all feasible costs and benefits to all sectors of society such as the provider, purchaser, clinician and patient.

Principles of economic analysis: costing

Opportunity cost

For an economic evaluation to provide decision makers with reliable and useful data it is imperative that the costing be carried out in a comprehensive manner consistent with accepted techniques. If this is not the case, the results will be invalid and unreliable hence severely affecting the credibility of the economic evaluation and in turn causing confusion over decisions around adoption of the intervention being considered. Only when the costs are carefully detailed alongside the benefits of an intervention can the decision maker decide whether the benefits are worth the costs involved. There is no point having excellent benefit data from a well-designed randomised controlled trial, say, when the cost data with which it is being combined are unreliable. For these reasons it is crucial that any economic evaluation of guideline implementation studies contain sound costing work. The following section is concerned with the costing principles. However, it is worth bearing in mind

that this is only half of the information required for an economic evaluation and that cost data are combined with benefit data, of which there is a section later in the chapter.

The economic concept of cost is "opportunity cost". This concept takes as its starting point the premise that resources are scarce. Therefore, every time we choose to use resources in one way, we are giving up the 'opportunity' of using them in other beneficial activities. The opportunity cost of developing and implementing guidelines is therefore defined as the benefit forgone from not using that resource in its best alternative use. Using this definition of cost, items to be included on the cost side of an economic evaluation are only those 'resources' which have an alternative use.

The importance of the margin

An important concept in costing (and benefit) exercises is that of the margin. The margin is concerned with change. The marginal cost is the cost of producing one more unit of a programme. Decisions concerning the allocation of scarce health care resources are usually concerned not with whether to introduce a service, but rather whether to expand or reduce a service. Given this, costing studies should be mainly concerned with measuring marginal costs. Jacobs and Baladi (1996) address the issue of bias in cost measurement each of which reflects the divergence of 'cost' from the desired 'marginal cost' measure, these are: scale bias; case-mix bias; methods bias and site selection. Site selection bias may occur when a cost, which is taken at one site, may misrepresent the marginal cost in the average site. This type of bias may be relevant for implementation studies, and care should be taken to establish costs at all the relevant implementation sites.

Discounting

Costs (and benefits) of health care interventions can occur at different times. For example, in guideline development and implementation programmes, as in prevention programmes, costs are incurred early in the scheme whereas the benefits may stretch years into the future. Individuals generally prefer to incur costs in the future (and receive benefits sooner). Given this preference, costs that are incurred in the future should be given less weight i.e. be discounted. Currently, the UK Treasury recommends a discount rate of 6% (HM Treasury 1982). For example, a guideline may be implemented which is costly to develop and implement at the outset but gives rise to a number of avoided inappropriate treatments and referral. These savings occurring in the future have to be discounted to reflect the fact that costs occurring in the

future impinge less on us that if they are incurred today. As a result, any beneficial savings arising in the future as a result of the guideline will have an 'apparently' lesser impact upon resources as compared to initial deve-lopment and implementation costs.

Sensitivity analysis

Every evaluation will contain some degree of uncertainty, imprecision or methodological controversy and as a result assumptions will have to be made (Drummond, Stoddard and Torrance 1987). For example, in implementing guidelines what if there were three outreach meetings instead of two? What would be the effect on costs if the guideline-recommended intervention gave rise to seven days in hospital instead of ten (the non-recommended but usual practice)? What if development costs were not included in the analysis because the local version was adopted from the national guideline and all that was being evaluated was the implementation? How long do the effects of the implementation strategies last and how long can the benefits be attributed to the implementation strategies? Sensitivity analysis allows the testing of the sensitivity of the results to the assumptions made. For a comprehensive summary of the main types of uncertainty and the corresponding role of sensitivity analysis in addressing this, see Briggs *et al.* (1994).

Categorising resources to be included within economic evaluations

Table 4.1 (page 84) provides some guidance on costs to be included in an economic evaluation of guidelines. In general, it is staffing costs which comprise the largest component of health care resources. With the implementation of clinical guidelines there may be a large component of the implementation costs which are due to staff time costs, hence it is crucial this cost reflect the true opportunity cost of that time. When reporting costs they should always be reported in the same year i.e. adjusting for the effects of inflation. The implementation of guidelines may also have an effect on other related services includes the staffing, supplies, overheads and capital costs associated with community, ambulance and voluntary services. Depending on the perspective of the study, costs to patients, their families and their friends may also require inclusion in an evaluation. Finally, indirect costs consist of time lost from work and costs external to health and welfare services (Donaldson & Shackley 1997).

Table 4.1. Costs and benefits of guideline introduction

Guideline Process	*Effect on whom?*	*Cost (+ve and –ve)*	*Benefit (+ve and -ve)*
Development (e.g. development meetings)	GP/Clinician	• Time and travel costs (£) • Opportunity cost of time and travel (work & leisure) • Resources e.g. secretarial costs	• Improvement in clinical knowledge • Job satisfaction • Problem solving skills
	Rest of NHS	• Resources e.g. consumables and printing costs	
Dissemination (e.g. outreach meetings and educational (workshops)	GP/Clinician/ Opinion leader/ Trainer	• Time and travel costs (£) • Opportunity cost of time and travel (work & leisure)	• Improvement in clinical knowledge • Job satisfaction • Problem solving skills
	Rest of NHS	• Resources e.g. consumables	
Implementation (e.g. structured medical records)	Patient	• Opportunity cost of patient time (work & leisure)	• Improvement in clinical knowledge • Job satisfaction • Problem solving skills
	GP/Clinician/ Opinion leader/ Trainer etc.	• Opportunity cost of GP/ Clinician/ Opinion leader etc.'s time e.g. time spent implementing the guidelines • Opportunity cost of time and resources e.g. change in number and length of consultations	
Intervention (e.g. change in clinical practice and resulting effects on resources and patients well-being due to guideline implementation)	Patient	• Opportunity cost of patient time (work & leisure)	PATIENT: • Health outcomes • Quality of life • Process benefits • Change in appropriateness of diagnosis and management • Patient satisfaction • Social outcomes GP/CLINICIAN/NHS: • Improved training of future GPs and clinicians • Improvement in clinical knowledge • Job satisfaction • Problem solving skills
	GP/Clinician	• Opportunity cost of GP/ Clinician time e.g. change in number and length of consultations	
	Rest of NHS	• Opportunity cost of time and resources e.g. change in number and length of consultations, waiting time, tests and procedures, number of clinics, inpatient stay, theatre time, staffing costs, emergency admissions etc.	
	Social services	• Opportunity cost of time and resources	
	Voluntary services	• Opportunity cost of time and resources	
Total - A summary of the total differences will be presented in +/- natural units of opportunity cost, +/- monetary units and +/- benefits			

Cost issues and guideline evaluation

When considering whether guidelines are 'worth' developing in the first instance, this involves prior consideration of the effect of implementing the guideline on *costs* as well as *benefits* of care. It is important to incorporate estimates of the cost of actually developing, disseminating and implementing guidelines along with the cost/benefit implications of any change in clinical practice achieved. There is no widely accepted successful way to incorporate cost considerations into guidelines. The Committee on Clinical Practice Guidelines (Institute of Medicine 1992) state that basic, accurate cost data are scarce and often not available and incorporation of cost data at all 'decision points' in a guideline may make that guideline unwieldy. These points are valid, but also apply to effectiveness data. Furthermore, they overlook the main problem, namely that of 'identifying the margins of change' i.e. those areas which may be expanding or contracting as a result of guideline implementation. Since marginal analysis looks only at the 'change', there is no need to cost elements common to the interventions, only those where a 'change' occurs.

Development costs

Development costs include: time costs of staff attending development meetings; time spent travelling to and from meetings; and research costs. There is also the issue of whether to develop a national or a local guideline. National guidelines may be relatively ineffective (Effective Health Care 1994); however, if the guidelines are aimed at a large number of clinicians (e.g. GPs) for a common condition (e.g. low back pain), then only a small uptake is required for them to be cost effective. However, if the uptake is relatively low due to a lack of local participation (i.e. no sense of 'ownership' (Maclean 1993)) it may be worthwhile, if anticipated incremental benefits are sufficient, to spend extra resources developing local guidelines. Attempts should be made to strike the most efficient balance between development costs and the dissemination, implementation and intervention costs. Table 4.1 above outlines some possible development costs.

Dissemination and implementation costs

According to 'Effective Health Care' (1994) educational interventions requiring more active participation by professionals, such as targeted seminars, educational outreach visits and the use of opinion leaders, are more likely to lead to changes in behaviour. Strategies for implementation include restructuring medical records and patient specific reminders. Strategies which are nearer the end user and integrated into the

process of care delivery seem to be more likely to be effective *(ibid.)*. Davis and Taylor-Vaisey (1997) show that interventions which are weak include mailings compared to those strong strategies such as reminder systems, academic detailing and multiple interventions. These strong interventions however are likely to be more resource intensive and hence whilst more effective are also more expensive.

All implementation strategies have associated costs that should be included within any evaluation, including: additional time spent by the clinician when using the guideline; patient time costs; and the hardware and software costs of computer assisted reminders. The potential costs incurred and benefits of implementation (and dissemination) are listed in Table 4.1. Finally, the table also outlines a section on '*Intervention*' with additional costs (or savings) and benefits due to changes in clinical practice as a result of guideline implementation. This latter cost/saving arises due to changes in management as a result of the guideline being effective in changing professional practice. This may legitimately increase costs where the guideline has suggested more tests, procedures or referrals or may decrease costs where the guideline has suggested a reduction in, say, inappropriate referrals or unnecessary tests and procedures. When the effectiveness of a guideline implementation strategy is being evaluated and data is already being collected on these process variables, a simple addition of the marginal cost to each of these items provides valuable economic evaluation information at the same time. Hence, whilst an economic evaluation is often thought of as a separate study, it is often the case that all the relevant economic variables are

Table 4.2. Incremental economic analysis based on 2x2x2 factorial design

Treatment group	*Cost (+ or -)*	*Benefit (+ or -)*
1. Control	Mean Baseline costs	Mean Baseline benefits
2. Educational intervention only	Additional mean cost over and above (1)	Additional mean benefit over and above (1)
3. Restructured medical record only	Additional mean cost over and above (2)	Additional mean benefit over and above (2)
4. Risk assessment only	Additional mean cost over and above (3)	Additional mean benefit over and above (3)
5. Educational intervention + restructured medical record	Additional mean cost over and above (4)	Additional mean benefit over and above (4)
6. Educational intervention + risk assessment	Additional mean cost over and above (5)	Additional mean benefit over and above (5)
7. Restructured medical record + risk assessment	Additional mean cost over and above (6)	Additional mean benefit over and above (6)
8. Education, medical record + risk	Additional mean cost over and above (7)	Additional mean benefit over and above (7)

inherent within the evaluation of effectiveness and all that is required is the addition of cost data. Table 4.2 outlines a hypothetical evaluation comparing the marginal costs and benefits of different implementation strategies.

Techniques of economic evaluation

Economic evaluation simply refers to the bringing together of the costs of the implementation strategies with the benefits of the strategies. In economic evaluation, the costing methods are all the same whichever technique is used, the only difference between each of the economic evaluation techniques are the way the benefits are valued. As noted earlier, only when costs and benefits are combined can a decision maker use the data to make a judgement about whether the benefits are worth the costs incurred. Hence whilst this chapter has dealt with costs and benefits separately, it must be remembered that they are equally important and the economic evaluation is only complete when they are brought together for a judgement to be made about a programmes' worthwhileness.

Benefit assessment

The inclusion of a health outcome measure is at the discretion of the evaluator, if the available evidence is sufficiently strong then all that is required is data on the effect the guideline has on the changing behaviour of the health care professional. This data on changed behaviour, along with the evidence on the effect of this on health outcomes can be modelled within the economic evaluation. Modelling is simply an exercise on which available data are combined (either manually or using computer spreadsheets) and the resulting outcome simulated rather than being observed. This is a powerful technique when the available data is evidence based, as is often the case for guidelines.

However, the method by which a guideline is implemented may induce health outcomes arising in a context which has not been quantified in previous literature, for example the health outcomes may have been measured and valued in an explanatory way and not in such a pragmatic implementation context. By including health outcome measures, one is by no means questioning any previous evidence merely quantifying it within a pragmatic implementation context. Clearly, the method by which evidence is implemented has varying success as a result of varying rates of uptake of the guideline. This is a result of varying rates of success of implementation strategies and the inclusion of a health outcome measure simply allows measurement of the success of implemen-

tation strategies. If one implementation strategy works better than another, this should directly impact upon health outcomes, hence the inclusion of such a measure will only serve to strengthen the evidence of the success of the implementation strategy. Such data, combined with data on the costs of the implementation strategy will provide rich data for the policy maker and aid in the decision of whether the guideline costs are justified by the benefits.

Richman, Scott and Kornberg (1998) included health outcome measures in their implementation evaluation of the outpatient component of an evidence-based disease management initiative. They measured the health status of children with asthma. In doing so, they can potentially provide high quality 'useful' cost-effective data for decision makers. Rush *et al* (1998) in their paper on consensus guidelines in the treatment of major depressive disorders note "whether guidelines actually improve outcome is largely uninvestigated, although a recent study of depressed patients in primary care found that using guidelines did improve outcome but at an increased treatment cost". They state that the clinical and economic impact of guideline-driven treatment for the severe and persistently depressed deserves study. Without inclusion of this outcome measure, the intervention would have been misleading as it would have appeared costly but at no extra benefit, which was not the case. Schell *et al* (1998) included measures of quality in their evaluation of the implementation of guidelines for the administration of perioperative antibiotics in bowel surgery. The results showed improvements in 4 of the quality indicators as well as modest cost savings.

Further to this, the implementation of a guideline may give rise to changes in the way a service is provided, in terms of the personnel involved, the process by which care is delivered, the setting in which care is delivered, the waiting time for care, the tests and procedures involved and so on. Hence, whilst there may be evidence on the actual intervention there may be little available evidence on patients valuation of these other process and non-health attributes. Health economics as a discipline has benefit assessment tools which are able to capture peoples' valuations of these changes and hence should be included within an implementation study where these effects are likely to occur. Exclusion of such valuations in an implementation study which has, say, streamlined the way a service is provided as opposed to providing major health changes, may underestimate the benefit of this to patients if such data are not collected.

The following section examines the techniques of economic evaluation. It is worth bearing in mind that whilst the costing exercise will probably be mandatory within any implementation economic evalua-

tion, it may be the case as outlined earlier, that there is in fact available benefit data which can be modelled alongside the cost data.

Economic evaluation

It is important to distinguish between costs and benefits within an economic evaluation. It is all too common for cost savings to be included as benefits when they are in fact negative costs (Donaldson & Shackley 1997). Benefits which may be included within a balance sheet of guideline implementation, other than health benefits where they occur may include: adherence to the guideline, improvement in clinical knowledge; appropriateness of diagnosis and management; patient satisfaction; and improvements in process of care.

The three principal economic evaluation techniques are: cost-effectiveness analysis (CEA); cost-utility analysis (CUA) and cost-benefit analysis (CBA). The technique(s) chosen will be determined by whether the question being addressed is concerned with allocative efficiency or technical efficiency. An allocative efficiency question is concerned with '*whether*' to allocate resources to a given programme. All health care programmes have to compete for scarce health care resources. These 'competing' health care programmes may include, for example, development and implementation of evidence-based guidelines for the prevention of deep vein thrombosis (DVT), the expansion of gynaecological services, the introduction of intensive care services and the development and implementation of guidelines for urology care. An allocative efficiency question would be: Should there be an expansion of surgery for hernia repair or should there be investment in the development and implementation of guidelines to improve the appropriateness of referrals in urology? In contrast, technical efficiency is concerned with 'within programme' efficiency i.e. '*how best*' to provide a given service. The resources, or budget allocated to a programme, are taken as given and the issue is simply 'how best' to provide that service. A technical efficiency question would be: When implementing evidence-based guidelines to prevent DVT, is it best to implement the guideline using restructured medical records or educational outreach meetings?

Cost-effectiveness analysis

Cost Effectiveness Analysis (CEA) is used to address questions of technical efficiency. It examines the effects of at least two competing alternatives 'within a fixed budget'. A ratio for each alternative is provided, the numerator being cost and the denominator the health effect. Such effects are measured in uni-dimensional terms i.e. life years saved or heart attacks prevented. The cost-effectiveness ratio produced is, there-

fore, a measure of 'cost per unit of effect'. The alternative with the lowest cost per unit of effect, or cost-effectiveness ratio, is the preferred choice. Hence, CEA is only useful if an implementation study is interested in only one outcome, e.g. % increase in appropriate referrals. The main limitation of CEA, not only in guideline implementation research but more generally, is that the unit of effect must be uni-dimensional. Hence, important effects may need to be excluded from analyses as a result of this. For example, in some implementation strategies, there may be important non-health outcomes and process outcomes which are valued by the patients but which could not be included in the uni-dimensional ratio. Another form of CEA is cost-minimisation analysis (CMA), this technique is used where the outcomes are identical and the comparison becomes one of costs only.

Cost-utility analysis

The benefit measure traditionally used in cost utility analysis (CUA) is the quality adjusted life year (QALY). QALYs were developed to take account of quality of life as well as quantity of life. To estimate QALYs, expected life years gained from given health care interventions are estimated and combined with information on the quality of these life years (via the estimation of utilities). CUA can be seen as an improvement on CEA as it attempts to combine more than one outcome measure, and takes account of both quality and quantity of life. The number of QALYs achieved from a health care budget will be maximised by allocating resources to those interventions with the lowest cost per QALY ratio. The use of CUA in evaluating guideline implementation would only be suitable where it was expected that there would be a resulting effect on the quality of life (QOL) of patients as a result of the implementation strategy. This may be the case where the guideline was not based on QOL data and information was required on the effect on QOL upon implementation.

Whilst CUA has become synonymous with QALYs, the technique can potentially be broadened to include measures of utility that take account of health outcomes, non-health outcomes and process attributes. Such an approach is appealing for implementation evaluators, as many potential changes due to implementation are of the non-health and process type. Using such an approach within a CUA framework, a cost per 'util' could be estimated rather than a cost per QALY. Utilities could be estimated using conjoint analysis (CA) (Ryan 1996). CA is a technique for establishing the relative importance of different attributes in the provision of a good or a service. These different attributes may be health, non-health and process attributes – all of which are likely to be impor-

tant in successful guideline implementation. For example, the successful implementation of a guideline to improve the appropriateness of urology referrals will not only change the number of appointments a urology patient will have to attend, but the management of the patient will change which may possibly mean more tests and investigations. This change in the management regime and changes in hospital stay, as well as possible changes in health outcome may be valued differently by the patient compared to their valuation of the traditional situation. Hence, there are many potential effects, which require inclusion if guideline implementation strategies are to be fully 'valued' in terms of health, non-health and process effects. This approach is in its infancy in health economics and further research is needed to look at how utility scores estimated from conjoint analysis studies can be used to address technical efficiency and allocative efficiency questions.

Cost-benefit analysis

Cost Benefit Analysis (CBA) is commonly used to address allocative efficiency, though it can also be used to address technical efficiency. Traditionally CBA requires all cost and benefits to be measured in commensurate units, usually money. Costs can then be directly compared with benefits. However, as a result of such monetary valuation of benefits, no ideal CBA has been carried out in the field of health care, despite the titles of many articles bearing the name (Zarnke, Levine & O'Brian 1997). Many cost benefit studies often turn out to be a comparison of costs incurred and savings accrued. However, this clearly involves only a comparison of costs with no consideration to the valuation of health benefits in monetary terms (Birch & Donaldson 1987).

Whilst there has been some progress in methods of monetary valuation of the benefits of health care, there has also been progress in using CBA as a framework for evaluation. This following section will examine the balance sheet approach, a proposed, practical prescription for the application of CBA.

Balance sheet

The balance sheet approach is a form of CBA which can be used to identify who bears the costs and who reaps the benefits from any change (McIntosh, Donaldson & Ryan 1999). Costs and benefits can be measured in physical units which seem both natural and appropriate. This approach adopts the definition of costs and benefits, outlined above, whereby: all effects on resource use are counted on the cost side, and all effects on patients' well being are counted on the benefit side (Birch & Donaldson 1987). Whilst the next stage in a CBA, as defined in health

economics, would require that all costs and benefits be valued in monetary terms, this is often not feasible or practical. The balance sheet approach, however, advocates that available monetary values can be augmented by other measures of cost and benefit – measures of quantity (e.g. numbers of referrals) and measures of time (e.g. time spent waiting for a consultation). This further highlights the role of CBA as an aid to decision making rather than as the sole criterion for those decisions (Sugden & Williams 1978); and it supports the view of Culyer, who, while recognising the imperfections of CBA in practice, also recognised its importance as a framework for decision making: "A good CBA will: identify relevant options for consideration; enumerate all costs and benefits to various relevant social groups; quantify as many as can be sensibly quantified; not assume the unquantified is unimportant; use discounting where relevant to derive present values; use sensitivity analysis to test the response of net benefits to changes in assumptions; and look at the distributive impact of the options" (Culyer 1985).

Gramlich (1997) also supports this approach to CBA, stating "Benefits and costs should be quantified when they can be and not when they cannot be, but whether quantified or not they should never be ignored. Even when they cannot be quantified, perhaps because they involve weighty matters of life and death, there are ways of setting up the analysis to focus public decisions properly". Whilst the balance sheet approach can be seen as a type of CBA in its own right, it can also be seen as the first stage in a CBA, i.e. as a means of outlining the benefits before monetary valuation. Whichever it is used for, it can be seen as a useful decision making framework in implementation research where the magnitude of the margins of change in terms of cost and benefit are still, in many respects, still unknown.

In Table 4.3, a hypothetical example of the balance sheet approach is presented. In this hypothetical example, the costs and benefits of developing and implementing a guideline to improve the appropriateness of urology referrals are presented. Included are costs in terms of guidelines development and implementation costs (e.g. staff costs and resources). Other cost implications in terms of preparation and freed-up clinic slots are reflected in their natural units because "monetary savings" would not necessarily reflect their true opportunity cost. Benefits are reported in terms of reduced waiting times, wellbeing improvements from reduced emergency admissions and reduced number of appointments, as well as improved satisfaction.

In this hypothetical example, it is clear to see how there would be problems with using a standardised quality of life questionnaire to quantify such a multitude of health, non-health and process effects on

Table 4. 3. Hypothetical example of the balance sheet approach in valuing the implementation of an initiative to reduce waiting time

Costs (+ve or –ve)	*Benefits (+ve or –ve)*
• Guideline development costs • Guideline implementation cost • Annual equivalent resource cost to develop and update guideline • 20 Hours *ad hoc* preparation, travel time and administration • Reduced emergency admissions by 20%, therefore cost savings • One less outpatient appointment to attend per patient (£41 per patient) • 300 Freed-up clinic slots (resource saving) • Patient time & travel cost savings of 50% (resource saving)	• Short waiting time of 1 week (compared to 7 weeks) • Increased wellbeing due to avoidance of emergency admissions by 20% • Increased wellbeing (or utility) from having to attend one less appointment • Increases wellbeing through earlier treatment and reassurance • Improved patient satisfaction

both patients and carers. Further, such quantification and, in a bid to adhere to convention, collapsing, of such effects into a cost-effectiveness ratio, may be done without consideration of any loss of useful information in this act. An economic evaluation can constitute a list of costs and benefits of alternative options and still provide useful information, arguably more useful than a cost-effectiveness ratio which is unable (by definition) to capture the many possible effects occurring. Often a description of the various effects on all the affected parties in a balance-sheet format may provide a more realistic and informative alternative to trying to estimate a cost-effectiveness ratio which, in fact, says nothing about the true effects or their diversity. At the very least, a balance sheet approach allows identification and measurement of the various effects on both costs and benefits occurring in guideline implementation. If this serves to identify one single effect, which then transforms the analysis into a CEA then the balance sheet has served a purpose.

CPP projects and CBA

Many of the CPP project teams did not have assistance from an economist. Hence, whilst many were keen to include economic considerations into their evaluations it was often not possible and not feasible to include economic considerations when the remit was to establish effectiveness first and foremost. However, it is in fact possible to identify the areas where the economics could have (and have in many instances) been slotted into each of the projects. The Danish project using computer assisted decision support (see Appendix, page 198) also has a variety of outcomes, which could benefit from a CBA approach. The primary

objective of this study is to evaluate the change of professional behaviour of the clinicians *and* the resulting quality improvement of the therapy. Parameters being measured, which could also be incorporated into an economic frame-work include: costs of development and implementation of the guidelines (e.g. costs of developing the computer software and capital costs); frequency of use of the computer system by clinicians and the opportunity cost of this time; duration of treatment and the cost; the frequency of patient information and the value of this information to patients; and clinical outcome i.e. relative duration of the patients' INR being within the therapeutic range. The CBA's balance sheet could incorporate the many potential costs and benefits of this study.

The Spanish CAMBIE project (page 216) may also find the CBA approach useful as a framework for an economic evaluation. The objective of this study is to produce and implement clinical practice guidelines aimed at reducing the unexplained variability and the rates of inappropriateness in patients with coronary heart disease. This study is estimating the change of physicians' behaviour after implementation of guidelines using feedback assessment of their clinical practice. The costs included in this study are the development and implementation of the guidelines. The cost of academic detailing, adaptation to local circumstances and feedback are all costs of the intervention. The costs in terms of opportunity cost may include staff time forgone, capital costs, time and travel costs, costs of academic detailing in terms of staff time forgone and resources, time spent during feedback and so on. These costs should be set aside the benefits, namely: rates of appropriateness and the resulting health gain or non-health gains of this reduced inappropriateness. Reduced inappropriateness may also have the effect of changing the management of patients in terms of reduced GP appointments, avoided outpatient appointment or even avoided hospital stay. Patients can place value on each of these possible beneficial effects which can be set aside the costs of the intervention and a decision made as to whether the cost of the intervention is *worth* the added gain in terms of health, non-health and process outcomes.

The Danish LysAMI project (page 195) is concerned with optimising the treatment of patients with myocardial infarction using the implementation of evidence-based clinical guidelines. Successful implementation of the guideline was expected to result in increased thrombolysis treatment rate and a quicker process of care in patients who met the treatment indications. The cost of developing and implementing the guidelines can be estimated as well as the costs/savings in the management of patients as a result of any changed professional practice. For example, increased thrombolysis treatment rates may give rise to avoided myo-

cardial infarction and hence give rise to savings in health care resources. Any savings in health care resources would have to be netted from the cost of developing and implementing the guidelines and set aside the benefits of the changed practice. In addition however there is also the added costs of the treatment being promoted in the guidelines which will have to be included as a cost, i.e. the cost of acetylsalicylic acid and thrombolytic treatment. The available evidence of the relationship between the treatment and the patient outcomes in this study means that changes in process measures can be used as reliable proxies for changes in health effect. As a result, it may be possible to directly estimate effects on health as a result of changed professional practice. In this situation, a CEA framework could be applied. However, since there are also changes in the process of care which patients may value, a conjoint analysis approach within a CBA may be more informative. With 16,000 Danes suffering from myocardial infarction, any change in professional practice which gives rise to treatment benefits and possible resource savings can be translated into a significant economic effect.

The Dutch CARPE (page 209) project is concerned with the implementation of guidelines for cardiovascular care with the aim of improving care for patients with cardiovascular risk indicators or diseases in general practice with trained facilitators. The central question in this study is whether and to what degree the care for patients with cardiovascular risk factors or diseases can be effectively improved by implementing this package of guidelines. This study has clearly outlined the costs and the benefits of the study. The possible effects of an improved quality of care as a result of the successful implementation of the guidelines on the quality of life and on the patient's satisfaction are being investigated. By combining the costs of the intervention with the effects on QOL (however measured) or satisfaction, a cost-effectiveness analysis of the guideline implementation strategy can be carried out. Given this, the additional costs and additional health effects arising due to the implementation of the guidelines can be estimated and a decision made as to whether these additional costs are worth the addition benefits.

Conclusion

There is no doubt that some types of guideline implementation strategies offer the potential to change professional practice and hence improve the effectiveness and efficiency of health care delivery. However, if they are designed without regard to efficiency there is a danger that they will lead to a net reduction in overall health in society. Throughout the CPP workshops it appeared that the main barrier to the economic

evaluation of guideline implementation strategies has been the perception of the complexity of the task. Hopefully now, this *perception* of complexity of identifying, measuring and valuing the costs and benefits within guideline introduction will be overcome, to some extent, by use of the 'balance sheet approach' as a framework for evaluation. However, by using the balance sheet approach as a starting point it may be that within this broad framework there transpires a clear case for using a narrower CEA or CUA approach. In this case, the balance sheet approach will have served a useful purpose. According to Williams (1994) 'Information on cost-effectiveness should influence clinical practice through practice guidelines (...) in principle there should be no resistance to this. After all, the objective of cost-effectiveness analysis is to ensure that the limited resources at our disposal are used to bring about the maximum improvement in people's health'.

KEY MESSAGES

- It is imperative that the opportunity costs of implementation strategies be identified, measured and valued
- To carry out a full economic evaluation, data are required on both the costs and the benefits of implementation research
- The various techniques of economic evaluation are useful for combining the costs and benefits of implementation strategies within a comprehensive framework
- The balance sheet approach is a pragmatic CBA method for identifying and measuring the costs and benefits of guideline implementation strategies

REFERENCES

Aucott JN, Pelecanos E, Dombrowski R, Fuehrer SM, Laich J, Aron DC. Implementation of local guidelines for cost-effective management of hypertension. A trial of the firm system. *Journal of General Internal Medicine* 1996; 11(3): 139-46.

Birch S, Donaldson C. Applications of cost-benefit analysis to health care: departures from welfare economic theory. *Journal of Health Economics* 1987; 6: 211-225.

Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. *Health Economics* 1994; 3: 95-104.

Culyer A J. *Economics*. Oxford: Blackwell, 1985.

Davis DA, Taylor-Vaisey A. Translating guidelines into practice: a systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines [Education]. *Canadian Medical Association Journal* 1997; 157(4): 408-416.

Donaldson C, Shackley P. Economic Evaluation. Pp. 849-71 in: Detels R, Holland WW, McEwen J, et al. (eds.). *Oxford textbook of public health. Vol.2: The methods of public health.* 3rd ed. Oxford: Oxford University Press, 1997.

Drummond MF, Stoddart GL, Torrance W. *Methods for the economic evaluation of health care programmes.* Oxford University Press, 1987.

Effective Health Care. Implementing Clinical Practice Guidelines: Can guidelines be used to improve clinical practice? *Effective Health Care,* December 1994, Number 8.

Gramlich E M. *A Guide to Benefit-Cost Analysis.* Second Edition. New Jersey: Prentice Hall, Englewood Cliffs, 1997.

HM Treasury. *Investment appraisal in the public sector.* London: HMSO, 1982.

Institute of Medicine. *Guidelines for clinical practice: from development to use.* (eds. Field MJ & Lohr KN). Washington: National Academy Press, 1992.

Jacobs P, Baladi JF. Biases in cost measurement for economic evaluation studies in health care. *Health Economics* 1996; 5: 525-529.

Lauterbach KW. Possibilities and limits of guidelines in medicine. *Zeitschrift für Ärtzliche Fortbildung und Qualitätssicherung* 1998; 92(2): 99-105.

Maclean D. *Clinical Guidelines: A report by a Working Group set up by the Clinical Resource and Audit Group,* May 1993.

McIntosh E, Donaldson C, Ryan M. Recent advances in the methods of cost-benefit analysis in health care: matching the art to the science. *Pharmacoeconomics* 1999; 15(4): 357-367.

O'Connor PJ, Solberg LI, Christianson J, Amundson G, Mosser G. Mechanism of action and impact of a cystitis clinical practice guideline on outcomes and costs of care in an HMO. *Joint Commission Journal on Quality Improvement* 1996; 22 (10): 673-82.

Richman MJ, Scott P, Kornberg A. Partnership for excellence in asthma care: evidence-based disease management. *Pediatric Annals* 1998; 27(9): 563-8.

Rosser WW, Palmer WH, Fowler G, Lamberts H, Thomson A, Lam C, Frame PS. An international perspective on the cholesterol debate. *Family Practice* 1993; 10(4): 431-8.

Rush AJ, Crismon ML, Toprac MG, Trivedi MH, Rago WV, Shon S, Altshuler KZ. Consensus guidelines in the treatment of major depressive disorder. *Journal of Clinical Psychiatry* 1998; 59 Suppl 20: 73-84.

Russell AS, Suarez-Almazor ME. Where's the evidence? [Letters] *Canadian Medical Association Journal* 1997; 157(12): 1736-1737.

Ryan M. *Using consumer preferences in health care decision making. The application of conjoint analysis.* Office of Health Economics, 1996.

Schell JA, Bynum CG, Fortune GJ, Laiben GR, Chang YJ, Pirner JA. Perioperative antibiotics in nonemergency bowel surgery: a quality improvement project. *Southern Medical Journal* 1998; 91(10): 900-8.

Sugden R, Williams A. *The Principles of Practical Cost-Benefit Analysis.* Oxford: Oxford University Press, 1978.

Vollman K, Sprung P, Posa S, Ladin D, Kachhal SK. Strategies for reducing material costs through implementation of clinical guidelines. *Journal of the Society for Health Systems* 1998; 5(4): 69-73.

Williams A. How should information on cost effectiveness influence clinical practice? Chapter 10 in: *Outcomes in Clinical Practice.* Tony Delamothe, BMJ Publishing Group 1994.

Zarnke KB, Levine MAH, O'Brien BJ. Cost-benefit analyses in the health care literature: don't judge a study by its label. *Journal of Clinical Epidemiology* 1997; 50(7): 813-822.

Chapter 5

Qualitative methods in implementation research

Nikki Rousseau, Elaine McColl, Martin Eccles, Lesley Hall

Aims of the chapter

This chapter aims to outline the role of qualitative methods within implementation research, and to suggest techniques which could be appropriate in specific situations. While a variety of potential applications for qualitative methods are discussed, the main examples are given from the COGENT (Computerised Guideline Evaluation in the NorTh of England) study, which involves qualitative research alongside a randomised controlled trial.

What is qualitative research?

"Qualitative research involves the collection, analysis and interpretation of data that are not easily reduced to numbers. These data relate to the social world and the concepts and behaviours of people within it." (Murphy *et al.* 1998).

Qualitative research is an attempt to present the social world, and perspectives on that world, in terms of the concepts, behaviours, perceptions and accounts of the people it is about. It is inductive; it generates theory from the data (as opposed to using data to test theory); it is holistic – it is concerned with context; and it appreciates that the social world is complex and tries to take this into account.

Qualitative research can be distinguished from quantitative research in terms of its aims and methods of data collection and analysis:

	Quantitative	***Qualitative***
Dominant paradigm	Positivist Natural science model Aims to find the general laws which will explain and predict behaviour	Antipositivist Does not believe that natural science model is appropriate for all research "Explanation by understanding" 1)
Research aims	To test hypotheses	To generate hypotheses
Research questions	How many? Which? Generally wants to fit individual cases into groups	What? Why? Seeks to understand what makes this individual case different
Research methods	Sampling – large, random, statistically representative Data collection – standardised Analysis takes place after data collection	Sampling – small, purposive Data collection – flexible Analysis and data collection happen simultaneously

Notes: 1) Bryman (1993)

An understanding of these differences is useful as it helps to explain the confusion amongst some quantitative researchers about the aims and methods of qualitative research. However, it is possible to place too much emphasis on these differences; in reality boundaries are blurred. Instead of arguing about the rights and wrongs of each approach, there needs to be an understanding of how the different methods can complement each other.

Why use qualitative research when studying interventions?

Qualitative research is often described as "naturalistic" – that it is useful for studying people in their natural settings. Why then would we want to use it when looking at the effectiveness of interventions? Randomised controlled trials (RCTs) are rightly considered the gold standard in evaluative research. The increasing use of the RCT to evaluate tech-nologies other than drug interventions has not been without difficulty; these problems have required the development of appropriate methodology (see Campbell *et al.* in this volume). When rigorously conducted, the RCT, which amongst other attributes, minimises bias and can prove causation, is an extremely helpful method of testing the effectiveness of a given intervention. For certain other types of questions, however (Why does this work? How does this work? Where do we go from here?), it is less appropriate.

This may be illustrated by reference to evaluations of the effectiveness of specialist stroke units. Early RCTs left many questions unanswered. For example: to what extent were good results simply due to the fact that stroke units were a new initiative staffed by highly motivated professionals? In practical terms a policy maker trying to apply the findings to another situation, would need to know what a stroke unit should consist of in terms of staff, size, facilities, etc. Understanding exactly how a stroke unit is different to standard stroke care and then identifying the elements leading to improved results is an extremely complex problem. A recent meta-analysis (Stroke Unit Trialists' Collaboration 1997) concluded that organised stroke unit care did lead to better outcomes for patients. However, while some of the more recent RCTs had collected information about the organisation of stroke units they had done so in a fairly ad hoc way. Thus, when they described the characteristics of stroke units which made them effective, the authors of the meta-analysis acknowledged that:

"several methodological problems exist with this approach to analysing stroke unit services. Firstly the information was obtained from the trialists who ran the stroke units and we were not able to obtain information from all staff who provided the conventional care. Therefore our findings could be biased by the expectations of the trialists as to which stroke unit features may or may not be effective. Secondly this was largely a retrospective analysis and in some cases specific questions could not be answered by the trialist or were not explicitly stated in the original published reports" (Stroke Unit Trialists' Collaboration, 1997).

Qualitative research could have helped people understand how exactly a stroke unit worked and how this was different from the conventional model of stroke care. It could have been used to identify what factors were particularly important to the stroke units' success and how care could be further improved. Having done this at an early stage in the testing of stroke units, later models could have optimised the stroke unit model or tested different models against each other.

Some possible questions for qualitative research:

How does this work?
What processes take place? Is there a smooth implementation or a catalyst which suddenly makes it take off? What is the catalyst?

Are we really seeing what we think we're seeing?
Is the copy of the guideline dog eared because of frequent usage or because its being used as a coffee mat? Is the GP really spending all that time interacting with the computer or is it on in the background while his consultation carries on regardless?

Why does (or doesn't) this work?
What aspects of this service do people like and what could be improved. Why? How could it be improved?

Where do we go from here?
How could the intervention be improved next time? Development of further research questions; development of hypotheses for testing quantitatively.

What (if any) have been the implications for this intervention beyond the immediate outcomes of quantitative interest?
Has the stroke unit led to similar innovations in other areas of the hospital? Or conversely has the siphoning off of enthusiastic staff and consultant time led to a reduction of morale in the general medical wards?

Some of these questions (e.g. which aspects work better than others) could be addressed by quantitative evaluation, but only after the initial domains have been established through qualitative work.

Relationship between qualitative and quantitative evaluation

Qualitative research can be used alone; it can also be used alongside quantitative research either before, during or after an intervention or implementation period.

Some uses of qualitative research when used alongside quantitative:

Before	Questionnaire development
	Optimising the intervention
During	Investigating processes
	Providing context
After	Explaining unexpected findings

Before a quantitative evaluation

It is now fairly widely accepted that the validity of a questionnaire is increased when it is based on an early phase of qualitative interviewing to establish the domains of interest and the language of choice of the population to be surveyed (Oppenheim 1992; Juniper *et al*. 1996). When questions have been devised, cognitive interviewing can help to ensure that they are user friendly and interpreted in the way intended.

As well as helping develop valid and user-friendly questionnaires, qualitative research could be used to optimise an intervention by identifying the requirements of the patients or health professionals who are to be the target of the intervention.

- Observational work and interviews could be carried out and documentary evidence (e.g. minutes of meetings) gathered to identify existing processes so that an intervention as far as possible works with rather than against peoples' existing systems of work. Similarly, qualitative methods could be used to design additional training to help people adapt to the requirements of the intervention.
- Important contextual information could be gathered; for instance, in general practice it might become apparent that administrative staff play a vital role in certain activities so that an intervention which took this into account would be likely to be more successful than one which did not.
- Strategic work could be carried out; group and individual interviews could be used to ask people what they would like; focus groups might be particularly good for suggesting new ways of doing things.

Qualitative research throughout an intervention or implementation study

Although quantitative methods *"can demonstrate that causal relationships exists, they are less useful in showing how causal processes work. Qualitative methods often allow the researcher to get inside the black box of experimental and survey designs and to discover the actual processes involved. Qualitative research is particularly good for*

developing explanations of the actual events and processes that can lead to specific outcomes. In this way, qualitative methods can yield theories and explanations of how and why processes and outcomes occur" (Kaplan and Maxwell 1994).

In clinical research, it is not enough to know that drug x works. We need to know in biochemical terms how and why it works; if we understand this, then we might be able to find something that will work even better, or to anticipate side effects. Similarly, in the social world, the use of opinion leaders can promote behaviour change (Thomson *et al.* 1998). However, to be able to replicate and optimise their use, we need to be clear how to identify opinion leaders and to understand why they have influence over others. In general, the more complex an intervention, the greater the need to understand the mechanism by which it works.

Qualitative research may be even more helpful where quantitative research can demonstrate a relationship but it is not clear whether the relationship is causal or the direction of the relationship. Some contextual information is commonly gathered in RCTs, and subgroups whose behaviour, experiences or outcomes differ from the main findings may be highlighted. However even where such data are gathered, quantitative findings have limited success in telling us why these groups are different.

Another important evaluative role for qualitative research refers to the intervention itself. Behavioural interventions are very complex – and it is possible to lose track of what is really going on. What was envisaged in the research plan may not be the intervention that finally takes place.

"The answer to why a program was ineffective may even reduce to the simple fact that it was not in reality operative: it existed only on paper" (Hyman *et al.* 1962, pp 74-5).

One study used video analysis of GP – patient consultations where computerised guidelines were used. This led to the identification of "verbal" and "physical" prescriptions; the GP first tells the patient that they are going to prescribe a certain drug (the verbal prescription) and then turns to the computer to actually produce a (physical) prescription. By that point it is too late for the computer to give advice; the prescription has already been given (Wilson 1999). So although the computer is in the consulting room, and the GP may even be using the system, he is not using it interactively in the way intended. Thus the intervention is not operative in the way envisaged. Hulscher *et al*. (chapter 7 in this volume) explore further the potential for qualitative research to assist in understanding the intervention.

Exploring unexpected results

While the ideal might be to conduct qualitative research alongside quantitative, thereby reducing the chance of obtaining unexpected results, this is not always possible. Much can still be gained in the event of unanticipated findings from conducting qualitative research with participants to explore why this might have occurred. For example, where it is possible for the intervention to remain in-situ, observational work could be done to identify any problems. Alternatively, focus groups could be convened with the targets of the intervention to learn about participants' experiences.

Triangulation

The use of qualitative and quantitative research together allows the possibility of triangulation. Triangulation refers to the integration of multiple data sources in research. The term stems from techniques in surveying or navigation whereby collecting more than one source of data leads to increasing confidence about your position (Campbell and Fiske 1969; Denzin 1970). Triangulation can also be "within method"; e.g. the use of observation and interviews in qualitative methods; the repeating of surveys at multiple time points in quantitative research. It is more powerful (and potentially harder to use) the more disparate the sources of the triangulated data.

"The effectiveness of triangulation rests on the premise that the weakness in each single method will be compensated by the counterbalancing strengths of another" (Jick 1983).

When applied to quantitative and qualitative data, triangulation can be used to validate results from each method but can also be used to generate new information, i.e. as an integral part of the analysis.

Techniques of qualitative research

Given the importance of well-conducted quantitative evaluations in implementation research, qualitative evaluations will often be taking place alongside these. This is the route we have chosen to take in the COGENT study. The remainder of this chapter therefore focuses particularly on how to use the methods together although much of the content will be applicable to qualitative research that stands alone. A brief introduction is given to the issues arising at each of the design, sampling, data collection, analysis, and reporting stages in qualitative research. There follows a discussion of the COGENT study, focusing in particular on design and sampling issues.

Research question

It is important to be clear about the research question as choices made at the design, sampling and data collection stages will affect the ability of the qualitative study to answer this. It is important to be clear at the outset whether the qualitative study is intended to be a stand alone study which is complementary to but entirely separate from any quantitative element, or whether it is primarily intended to illuminate the findings of the quantitative study. The implications of this decision are discussed in the section on sampling in the COGENT study.

Design

Quantitative research is usually classified in terms of its overall study design (randomised controlled trial, cohort study). By contrast, qualitative research has tended to be categorised in terms of its methods of data collection (focus groups, observation; e.g. Fitzpatrick and Boulton 1994). Perhaps because of a reluctance to "fix" the study at too early a stage, qualitative researchers tend not to speak of design in a way that is familiar to quantitative researchers. Yet design is still very important to qualitative research. In applied health research it is seldom either possible or desirable to allow qualitative research to be a truly organic process. Therefore it is advisable to give some thought early on in a qualitative study to such questions as:

- do I want to follow people over a period of time or will a single snapshot be enough?
- is the organisation an important part of my research question; do I need to consider interviewing a number of people within the same organisation, or would it be better to look at people from a range of different organisations?
- do I want a "control" group; do I want to understand how this is different from other systems, or am I primarily concerned with understanding *this* system?

Sample

Sampling in qualitative research is described as "purposeful". The aim of the sample is not to cover a statistically representative part of the population, but to ensure that the range of relevant behaviours and/or attitudes relevant to the research question is covered (Mays and Pope 1995). In the grounded theory approach (Glaser and Strauss 1967) the sample is not pre-set at the start of the research. Instead, sampling develops in response to theory which itself develops as data is collected and continues until the researchers are confident that no new or conflicting evidence remains unidentified. Researchers deliberately seek out the cases that do not appear to fit the patterns being uncovered, rather

than aiming to make generalisable statements about the sample studied. The use of a sampling frame (an example is given in the description of COGENT below) can help to ensure that sampling is systematic and transparent. As with a quantitative study this sampling frame should be based on the research questions and take into account what is already known or hypothesised.

Data sources and collection

Many techniques of data collection are available to the qualitative researcher:

Participant observation	Researcher takes an active role in the organisation being studied. Researcher can trouble shoot problems, adapting intervention to suit case. This can be very useful if the aim is to optimise the intervention.
Non-participant observation	For example, of practice meetings or patient consultations. Whilst the impact of a researcher cannot be totally eliminated, the researcher observes without offering advice or attempting to alter the course of events.
Video	Variation on the above, particularly suitable if detailed analysis of interaction (e.g. patient-doctor) and conversation are required.
Individual interviews	Of varying structure and depth. In "semi structured" interviews, the researcher will have a fixed list of topics to cover. In unstructured interviews only a very broad outline of the area interest will guide the interview; the interviewees priorities will determine exactly what is covered. Both what people say and how they say it may be of interest. Qualitative interviewing is highly skilled work and needs to be carried out by researchers who are very clear about the objectives of the study.
Focus groups	Useful for obtaining views from a number of people in a short space of time. Group processes are inevitable; when studying a group of people (as in a GP practice) this can be a bonus (for example in uncovering important details about relationships within the organisation or because people within the organisation will challenge their colleagues over a discrepancy between what they claim to believe and what they actually do (Kitzinger 1995)).
Documents	Minutes of meetings, annual reports or pre-existing guidelines can be valuable sources of data.

The methods of recording data are important in qualitative research and generally need to be as full as possible (to enable researchers to return to the "raw data" when necessary). Audio and video recording can be invaluable, particularly in interviews – where it is generally impossible to

record data at the same time as generating it by relevant questioning. The tapes are then transcribed; sometimes by the researcher as part of the initial process of (re)familiarisation with the interview.

Analysis

Just as sampling in qualitative research is not carried out solely at the start of a project, so analysis is not carried out solely at the end of a qualitative study. Ideally, analysis should commence as soon as the first data has been collected, to enable the sample and data collection to be developed and refined. This also allows the researcher to know when "data saturation", the point at which new data collection is no longer adding anything to the analysis, has been reached.

Depending on the aim of the research different types of analysis can be carried out, e.g. looking not only at content (e.g. the consultation is about treatment for asthma) but who talks (e.g. doctor dominates the conversation) and the type of interaction (e.g. doctor is being directive).

In practical terms, qualitative analysis consists of a process of reading through transcripts and other data, identifying themes, coding data (applying these themes to data) and then drawing together themes and cases. This process can be done manually (the "Framework" method [Ritchie and Spencer 1994] is popular) or with the help of a computer. There is sometimes confusion about the role of a computer in qualitative analysis; it is primarily a tool for helping the researcher manage the large and complex dataset and does not replace to researcher in the task of reading through transcripts and identifying and applying themes. Whether the task is done manually or with a computer it must be done systematically and not simply by reading through notes until a quote is found to support a particular theory.

Reporting the results of a qualitative study

When presenting the results of qualitative study it is important to:

- Make transparent the process of data collection and analysis – the reader should know enough about the methods used to be able to repeat the process if they wished to.
- Present enough data for the reader to be able to judge whether the interpretations made by the authors are fair.
- Present enough information about the cases presented for the reader to be able to relate the cases presented to other situations.

Example: Design and sampling issues in the qualitative element of the COGENT study

The COGENT case study

Within COGENT (see box below), the RCT will tell us the effect of computerised guidelines on process and outcome of care but will tell us little about how or why change is brought about. This is the area that the case study will allow us to explore. These questions are vitally important to the deve-lopment of knowledge and technology in this area and will make the findings of the RCT much more useful to researchers and policy makers. Thus we might expect the case study to examine whether there are prac-tical problems with the implementation of the guidelines, and why any observed variation in use exists.

The COGENT study

COGENT is the evaluation of computerised guidelines for the management of two common chronic conditions managed in primary care: asthma and stable angina. The study is funded by the UK National Health Service National Research & Development Programme and is run by a collaborative trials group from a number of university departments. The study comprises a randomised controlled trial (evaluating impact of guidelines on process and outcome of care using secondary data and before and after surveys of patients), an economic evaluation and a case study. The intervention is the computerised delivery of the North of England evidence-based guide-lines for asthma and angina (North of England Evidence Based Guideline Development Project 1996a and 1996b). GP practices from the north of England were invited to participate if they were using a MEDITEL or EMIS computer system, and had a majority of partners using the computer interactively during GP-patient consultations. Each practice will receive a computerised version of one of the guidelines and a paper version of both. If practices already have guidelines for one or both of the study conditions, the process of study guideline implementation may allow a degree of customisation in line with their current practice, to encourage ownership of the

Design

In the COGENT study we chose a case study design for the qualitative aspect to the evaluation. A case study is a research design that involves a very detailed study of an individual or organisation (the "case"). Multiple sources of data are commonly used to draw up a comprehensive

picture of the case. Thus a case study will include various types of qualitative data collection; observational work, interviews etc., and often also incorporate some quantitative data (though in a qualitative case study, analysis is predominantly qualitative in nature). A case study design was considered most appropriate for two main reasons:

Implementation is a process
Implementation is not a one-off event but rather a process, with several stages which take place over time. We wanted to be able to understand these stages. Unless we investigated the implementation over the whole period there was a danger we might miss something important.

Health professionals do not act in isolation
This is one reason why case studies are becoming popular in health services research. While behaviour change ultimately has to take place at the level of the individual, interventions often take place at the practice level; indeed it may be difficult to stop an individual level intervention spreading to the practice. Similarly social and organisation factors beyond the individual will affect the success of an intervention. The individual health professional will be acted on by various forces both within and outside the practice; similarly they will exert forces on their environment. In quantitative research this is the rationale for cluster randomisation as Campbell *et al*. have discussed in chapter 3 of this volume.

Whilst much variation in the use of computerised guidelines appears to be at the level of the individual, the social and organisational setting (the most important aspect of which is the practice) provides important context for the intervention. The use of case studies will allow us to gather information about the context of the evaluation in a way which would not be explicit in a design where the focus was on individual GPs selected from various different practices. The practice itself is a case, and the individual health professionals within the practice are also cases. This type of design with both practice and individuals as units of analysis is known as an **embedded design** (Yin 1984).

For practical reasons (data collection needs meant post-hoc exchange of practices between the RCT and the qualitative element was not possible; resources did not allow for recruitment and training of additional practices) it was not possible in COGENT to develop the sample of practices once the intervention period had started. This also meant that the sample of individuals would be limited to those in the practices we chose.

Our first decision was whether to choose, as cases, practices within or outside the RCT. This decision is related to the research question –

whether understanding what happens in the research processes surrounding the RCT (rather than purely the intervention itself) is of interest. We had various approaches available to us for the selection of case study practices:

Possible sampling strategies where qualitative and quantitative research is conducted together:

Approach	*Pros*	*Cons*
Use units taking part in quantitative study (RCT practices in COGENT)	This is the only option if no additional units available. Ensures any implications of research activities surrounding the quantitative study are picked up (e.g. patient activity in response to questionnaires prompts action)	Risk of contamination of RCT e.g. because interviewer acts as a reminder to use the guidelines, overloading practices with the data collection requirements of both methods
Use units eligible[1)] and willing to participate in quantitative study but not included (i.e. part of sampling frame for quantitative study)	Allows implementation in qualitative research practices to be closer to "real life"; not contaminated by survey activities	Additional practices required. May not pick up issues to do with research activities which are specific to the quantitative research and which might affect the success of the implementation (e.g. in COGENT RCT practices had to run time consuming computer searches which may have led to negative feelings towards the study)
Use units willing to participate but not eligible for quantitative study	Answers the question "what would be the impact of the intervention in practices which are different from those in the quantitative study" (in COGENT – because they were already using a version of the guidelines or because they were not using the computer interactively)	Resources needed for extra practices, or reduces resources available to study main research question
Use practices not previously recruited for study RCT	As above; tackles the question about implementation in different situations	In COGENT we were restricted to practices using certain computer suppliers because of the huge software costs. We had already approached all those practices within easy reach of the research team

1) Practices eligible for participation in the RCT are those with an EMIS or Meditel computer system, those with more than one partner (to ensure sufficient patients for the patient outcome survey) and those where over 50% of partners are using their computers "interactively" (for viewing clinical data during a consultation and for acute prescribing).

In COGENT we went for the second approach; to use practices eligible and willing to take part in the RCT, but which would not be included. At this stage we had not preselected either the RCT or case study

practices so we needed a method of choosing the case study practices from the total pool of practices available to us.

Whilst it was important when developing the sampling strategy to take into account what was already known (for example about individual and organisational behaviour and human – computer interaction), we wanted to remain open to other possibilities. We therefore decided to include a random element to the sampling strategy. We allocated all the practices situated too far from Newcastle for it to be feasible to include them in the case study, to the RCT. We then randomised the remaining practices to be in the RCT or in the pool for selection for the case study. We then selected case study practices from the pool on a purposeful basis (if there had not been practices in the pool which met our criteria we would have taken practices from the RCT group – however this did not prove necessary).

The following features have been hypothesised to influence the successful implementation of an innovation

Features of the individual GP	Attitude to guidelines, computers and the research study and team
	Current use of computers and guidelines
Features of the organisation	Presence of a strong advocate for the guidelines and/ or for the study
	Presence of enabling administrative or structural features
	Type of computer system
Features of the innovation	Ease of use
	Appropriateness
	Compatibility with existing practice (size of change required)

These features informed our choice of practices, though we did not always have data which allowed us to use the criteria directly.

COGENT General practice sampling frame

Criteria	*Theoretical basis*	*Sampling strategy*
Similar to RCT practices	Primary purpose of case study is to illuminate findings of RCT, therefore it is important that practices are not too dissimilar to RCT practices in the following areas:	Broadly similar proportions to the RCT practices in the areas outlined below
Computer system	We thought that the implementation of computerised guidelines might prove to be technically neater on one or other system or that there might be some other effects to do with the computer (training and back up, self selection effects) which would enhance implementation	At least two of each computer system in our sample of 6
Fundholding wave	There is some evidence that fundholding practices might be more responsive to change and positively inclined towards innovations than other general practices. There might be additional features associated with fundholding that would encourage practices to use the computerised guidelines	At least one non fundholding practice and at least one phase one fundholding practice
Training practice	As above. Vocational training practices should have a positive attitude to teaching and research and will have to meet certain organisational standards.	Not more than half of our sample were training practices
Number of partners	We felt that practice organisation and communication between partners might differ between small and large practices and that this might affect the uptake of the guidelines	A mixture of larger and smaller practices
Part time partners	We felt that part time partners might find it more difficult to adopt the changes.	Some practices having part time partners
Already have guidelines for asthma and or angina	These practices might be more positively inclined towards guidelines. Alternatively, differences between the two guidelines might be a disincentive to use the computerised guidelines	A mixture of practices having and not having guidelines
Have some partners not using computer interactively	We were interested to see the impact of a computerised guideline system on those partners not currently making as much use of computers as their colleagues	Some (but not all) practices having some partners not making interactive use of the computer

Sampling of individuals

At the individual practitioner level, sampling in COGENT is using several approaches:

Pre-intervention

- All contact persons (the person the practice identified as "best to contact about COGENT issues") within practice
- People identified as useful data sources during a preliminary interview with contact person
- Results of an attitudes and experience questionnaire completed by GPs and practice nurses (including both typical and extreme cases)
- Other people who become relevant as theories develop (e.g. interviews with health professionals may lead us to hypothesise that practice managers are key and they could then be interviewed).

Post-intervention

- Follow up of original cases
- People mentioned during interviews
- Computer use log files (recording amount and type of interactions with the computerised guidelines) to identify new cases (as well as informing discussion with existing cases being followed).

Post RCT

- Individuals from the RCT could be used to explore whether experiences in the case study practices were shared in the RCT sites.

Patients

Patients are clearly a crucial component in health services research. For the purposes of the COGENT case studies we decided that our primary focus was the health professional; we might observe patients in the consultation but we would not interview them directly. Where implementation research involves patients more directly (e.g. patient held guidelines), it would be desirable to include interviews with them.

Key messages

Most implementation evaluations would benefit from the inclusion, at appropriate points, of qualitative approaches. Good qualitative research is not easy, and careful thought needs to be given to each of the stages of design, sampling, data collec-tion, analysis and reporting of results.

Further reading

The following texts may be useful for those wishing to use qualitative methods within implementation research:
Patton MQ 1987: How to use qualitative methods in evaluation.
Yin RK 1984: Case study research: Design and methods.
For exact references, see below.

References

Bryman A. *Quantity and quality in social research.* London: Routledge, 1993.

Campbell D, Fiske D. Convergent and discriminant validation by multitrait multidimensional matrix. *Psychol Bull* 1959; 56: 81-105.

Denzin NK. *The research act in sociology: a theoretical introduction to sociological methods.*London: Butterworths, 1970.

Fitzpatrick R, Boulton M. Qualitative methods for assessing health care. *Quality in Health Care* 1994; 3: 107-113.

Glaser B, Strauss A. *The discovery of grounded theory.* Chicago: Aldine Publishing, 1967.

Hyman HH, Wright CR, Hopkins TK. *Applications of methods of evaluation.* Berkeley & Los Angeles: University of California Press, 1962.

Jick H. Mixing qualitative and quantitative methods: triangulation in action. In: Van Maanen J (Ed.).*Qualitative methodology.* Beverley Hills:SAGE Publications. Inc., 1983.

Juniper EF, Guyatt GH, Jaeschke R. How to develop and validate a new health-related Quality of Life Instrument. Pp. 49-55 in: Spiker B (Ed.) *Quality of Life and pharmacoeconomics in clinical trials.*2nd edn. Philadelphia: Lippincott-Raven, 1996.

Kaplan B, Maxwell JA. Qualitative research methods for evaluating computer information systems. Pp. 45-68 in: Anderson JG, Aydin CE, Jay SJ (Eds.) *Evaluating health care informations: methods and applications.* London: SAGE Publications, 1994.

Kitzinger J. Introducing focus groups. *British Medical Journal* 1995; 311: 299-302

Mays N, Pope C. Rigour and qualitative research. *British Medical Journal* 1995; 311: 109-111.

Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P. *Qualitative research methods in health technology assessment: a review of the literature.* 2. The National Coordinating Centre for Health Technology Assessment. 1998

North of England Evidence Based Guideline Development Project. *Evidence based clinical practice guideline: the primary care management of stable angina.* Report 74, Newcastle upon Tyne: Centre for Health Services Research. 1. 1996a.

North of England Evidence Based Guideline Development Project. *Evidence based clinical practice guideline: the primary care management of asthma in adults.* Report 75, Newcastle upon Tyne: Centre for Health Services Research. 1. 1996b.

Oppenheim AN. *Questionnaire design, interviewing and attitude measurement,* 2nd ed. London: Pinter Publishers Ltd., 1992.

Patton MQ. *How to use qualitative methods in evaluation.* 2nd ed. Newbury Park, California: SAGE Publications, Inc., 1987.

Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RG (Eds.) *Analyzing qualitative data.* London: Routledge, 1994.

Stroke Unit Trialists' Collaboration. Collaborative systematic review of the randomised trials of organised inpatient (stroke unit) care after stroke. *BMJ*, 1997; 314: 1151-1159.

Thomson MA, Oxman AD, Haynes RB, Davis DA, Freemantle N, Harvey EL. Local opinion leaders to improve health professional practice and health care outcomes. In: Bero L, Grilli R, Grimshaw J, Oxman A (Eds.) *Collaboration on effective professional practice module of the Cochrane Database of systematic reviews (updated 01 December 1997).* Available in the Cochrane Library (database on disk and CDROM). The Cochrane Collaboration; Issue 1. Oxford: Update Software, 1998.

Wilson R. *Verbal and physical prescriptions in general practice consultations.* 52, Newcastle upon Tyne: Sowerby Unit for Health Informatics, 1999.

Yin RK. *Case study research: design and methods,* 5th ed. Beverly Hills, California: SAGE Publications, Inc., 1984.

Section C

Planning and monitoring the implementation Addressing barriers and facilitators

Chapter 6

Methods for identifying barriers and facilitators for implementation

Michel Wensing, Miranda Laurant, Marlies Hulscher, Richard Grol

Aims of the chapter

This chapter describes and discusses a range of observational research methods that can be used for identifying the barriers and facilitators for changing professional practice in health care.

Introduction

Successful implementation of clinical guidelines and research evidence into practice is determined by a wide range of factors. For example, Dutch general practitioners (GPs) who were asked to implement guidelines on cholesterol had doubts about the cost-effectiveness of cholesterol screening for some of the patient categories described in the guideline; preventive services did not have a high priority for the GPs who were reluctant to interfere with patients' life-style; providing these services was time consuming, as three cholesterol measurements were needed according to the guidelines (Van der Weijden 1998). Such concerns are of course legitimate, but it is desirable to know them well before implementing a guideline. These are examples of typical *barriers* that impede the adoption of clinical guidelines. There may also be identifiable *facilitators* that enhance the implementation process. For example, support to a guideline from a well-known and well-reputed local professional may encourage others to apply the guideline in their clinical work.

In more general terms, factors that influence change of clinical practice may be related to the knowledge, skills and attitudes of the clinicians or to the social, organisational, economic and legal context in which they work (Lomas 1996). Insight into such factors is important because implementation strategies that take into account the relevant barriers and facilitators for change are expected to be most effective (Grol 1997); and this is important also because there is evidence that the effect of implementation strategies is usually modest at best (Grimshaw 1995, Oxman 1994, Wensing 1998).

When barriers and facilitators for change are measured before an implementation strategy is chosen and applied, an intervention can be tailored to the targeted clinicians and their work setting. Such tailoring based on explicitly identified barriers and facilitators is not yet common: a systematic review of studies on continuing medical education showed that this was done in only 18% of the educational interventions studied (Davis 1995). Empirical evidence on the effectiveness of tailored strategies based on barriers/facilitator analysis is scarce, even if most theories on behaviour change would suggest that such an analysis should be made. Educational theories suggest that learning needs should be assessed with gaps analyses; behaviouristic theories talk about incentives and sanctions; and management theories are interested in describing inadequate fits within or between systems.

Identification of barriers and facilitators *during* or *after* an implementation process may help us understand why a guideline or an intervention was effective in some practices, but not in others. For instance, the effect of any intervention to improve cholesterol management may be higher in larger practices, because there is more supporting staff than in small practices, and this might be discovered through targeted studies.

This chapter addresses how barriers and facilitators can be studied. But as the experience in this particular field (guidelines implementation) is scarce, the following is an overview and discussion of seemingly relevant methods, more than a systematic review of evidence-based strengths and weaknesses of each method.

Materials and methods

The overview of observational methods is based on: (a) a secondary analysis of a review of implementation studies (Wensing 1998), focused on the identification of barriers and facilitators for change and on the methods that were used for that; (b) computerised searches in Medline and Psychlit (period 1980-1997), focused on specific methods that were thought to be useful, followed by snowball sampling in the publications found; (c) suggestions and examples from researchers working on im-

plementation studies. We considered all studies on the dissemination and implementation of innovative procedures or technologies in clinical practice: our scope was broader than clinical guidelines, because we did not want to exclude relevant studies.

Barriers and facilitators

Barriers and facilitators are defined as factors that actually prevent or enhance, respectively, changes in clinical behaviour. For a global classification of barriers and facilitators for change we refer to the CPP framework described in chapter 2, which can be used as a checklist for considering potentially relevant factors. This chapter focuses on characteristics of the clinicians and their work setting that can be observed to influence implementation. We do not discuss planned interventions to implement guidelines (such as education or feedback) or characteristics of the guidelines (such as their format or specificity), although these also may influence guideline uptake.

Variations in clinical behaviour, such as number of tests ordered, surgical procedures performed or adherence to a specific guideline may be related to barriers or facilitators for change. For instance, test ordering is related to clinicians' attitudes toward uncertainty, difficulties in using the rational decision-analytic approach, or perceived patients preferences (Grol 1996). Insight into these factors can be gained from studies on variation of clinical practice. Other factors are related to the change process itself; these emerge in the process of changing routines, but are difficult to predict beforehand. For instance, a clinician may enjoy discussions about clinical guidelines with colleagues, or experience unforeseen financial consequences after improving the practice organisation to meet guideline recommendations.

Perceptions and opinions that clinicians have of barriers or facilitators for change may be adequate or inadequate. They are relevant if they influence changes in behaviour: the perceptions may be unrealistic, but real in their consequences. On the other hand, factors that actually influence behaviour change may not be perceived by clinicians. A well known example of this is shown by a study where CME was only effective in clinical areas that had low priority for clinicians (Sibley 1982).

A crucial consideration is whether factors that influence behaviour change can be manipulated. For instance, age of the clinician may be a predictor for test ordering, but it cannot be changed. However, age may be a proxy measure for specific beliefs and attitudes which might be changeable. Identification of barriers and facilitators with the aim to design or modify implementation strategies should focus on factors that can be changed.

Routine data/databases

In most health care systems data on clinical procedures (referrals, prescriptions, test ordering, etc.) are collected routinely. Such data may also be collected for a specific study. In both cases the data can be used to hypothesise – or at best: determine – which factors influence clinical decision making as barriers and facilitators. Ideally, this analysis starts with the development of explanatory models which are then tested on the available data. For instance, between 50% and 90% of decision making on prescriptions could be predicted with decision making models (Segal 1985, Chinburapa 1987, Denig 1988). The use of data on clinical procedures is probably most useful to verify the relevance of different factors on a large scale, but its feasibility depends on the availability of the data and skills to analyse these adequately. An example of the use of routine data is given in Box 6.1.

Box 6.1
Health care utilisation for low back pain (Rossignol 1996)
A prospective cohort study (n=2147 patients) was performed in Quebec, Canada, to analyse use of health care in relation to clinical guidelines on low back pain. The guidelines advice to use imaging tests in absence of alerting symptoms only after 1 month and when considering surgery. 4.5% of the patients received these tests at least once, often with a delay of 2-5 months. The study showed more testing for patients with specific diagnoses (OR=4.5), for those living in rural areas (OR=2.0) and in primary occupations (OR=2.4).

Survey methods

Written questionnaires for clinicians have often been used to identify perceived barriers and facilitators for implementation of guidelines or innovations in health care. An advantage is that large numbers of respondents can be included in a study at a reasonable cost. The problem is finding out the extent to which these perceptions or attitudes are related to actual behaviour. Psychological research suggests that often there is only a moderate relationship between attitudes and behaviour, but this relationship may be different for perceptions. We believe that the identification of barriers and facilitators should not be based on a survey alone, but that a survey may be useful to generalise factors identified in an in-depth study.

Questionnaires on guidelines

A large number of questionnaires on guidelines (attitudes, uptake etc.) has been developed (see chapter 9). These questionnaires differ in focus, type of questions and answering format.

Foci may be clinical guidelines in general, a limited set of guidelines, a specific guideline, or specific recommendations within a guideline.

Types of questions may be recall or knowledge of the guideline; opinions on the guideline; perceived barriers and facilitators for implementation.

Answering format can be for instance 5-point Likert scale (agree-disagree), yes/no questions, other formats.

See example in box 6.2.

Box 6.2

Newton et al. (1996) in England used a questionnaire sent to 300 GPs on attitudes towards clinical guidelines:

Focus

3 guidelines:
- Asthma
- Radiology
- Diabetes

Types of questions

1. Questions on knowledge/recall of guidelines
2. Questions on attitudes regarding guidelines
3. Questions on barriers and facilitators for implementation

Answering format respectively

1. 3 items on knowledge of the content of the guidelines (5-point scale, 'never heard of' to 'very familiar with')
2. 13 items on guidelines in general (5-point agree-disagree). Example: 'narrow clinical freedom'
3. 7 items on feeling that a range of persons or agencies pressure to use guidelines (5-point scale: 'strong pressure' – 'no pressure at all'). Example: MAAG. 8 items on likely impact of methods used to facilitate the uptake of guidelines (5-point scales, 'very likely to make me use the guidelines' to 'not at all likely'). Example: Reminders from source.

Our empirical basis does not make strong evidence-based recommendations possible, but our experience with interviewing clinicians on guidelines suggests that the questions should focus on a specific guideline or a limited set of guidelines.

Questionnaires on specific cases

The problem of the weak relationship between perceptions and actual behaviour may be overcome to some extent by filling a questionnaire shortly after applying a guideline. Experience with this method is limited and it may be too time consuming (Box 6.3). Nevertheless, we believe that this method deserves further attention in future research.

Box 6.3

A questionnaire on cases of defensive behaviour (Van Boven 1997)

Defensive behaviour is a clear deviation from usual behaviour and good practice in order to prevent complaints or criticisms by the patient or her family. Over a one-year period 16 family physicians with 31343 patients recorded all episodes of care involving an order for laboratory tests, diagnostic imaging, or both (n=8897). The physicians selected one or more reasons to order each test from a fixed list of clinical considerations. In addition, they recorded whether they acted defensively for every test order.

Questionnaires on potential determinants

Numerous potential determinants of the variation in clinical behaviour and behaviour change have been derived from various theories. Examples include use of information sources (communication and innovation diffusion theories); learning style and learning needs (adult learning theory); self-efficacy – the perception that a task can be performed (social learning theory); and opinion leaders for the targeted clinicians (innovation diffusion theories). Written tests of knowledge can also be included in this category as some insights or skills are needed for most changes (cognitive theories). Written questionnaires (example Box 6.4) have been developed and (sometimes) validated to measure such factors, and some of these are supported by empirical evidence.

The usefulness of questionnaires depends on the actual relevance of the studied factor for change in clinical behaviour. For instance, the relationship between competence (as assessed by a test of knowledge) and performance is moderate at best (Rethans 1996). Usefulness also depends on the validity, reliability and feasibility of the questionnaire it-

BOX 6.4
AN EXAMPLE ON SOURCES OF INFORMATION (Dolan 1997)
Sample: 46 physicians on one hospital, USA.

Focus: Published decision analysis

Types of question and **answering format**

- Questions on awareness and understanding. Format: 3 items on understanding, training and interest regarding decision analysis (answering scale unknown)

- Questions on actual use and experience. Format: 1 item on use of a published decision analysis to guide management of a patient (answering scale unknown)

- Questions on attitudes and preferences. Format: 13 items on helpfulness of 13 interventions to improve use of decision analyses (4-point scale, not helpful at all big help). Example: Easy access to latest review of the topic from your office. 3 questions on level of interest in educational activities to learn about decision analyses. Example: on-time educational session.

self, which is difficult to separate from the factor being studied. These questionnaires cannot be recommended for practical planning of implementation strategies before further research on the relevance of different factors is available.

INTERVIEW METHODS

Oral interviews with clinicians, usually face-to-face but sometimes by telephone, have been used to identify perceptions and opinions regarding guidelines or other innovations (see Box 6.5). Although the problem is again to what extent actual behaviour is related to clinicians' answers to the question, the advantage compared to questionnaires is that the interviewer can ask for clarification and underlying motivation. The method is time consuming, and often only a small number of clinicians can be interviewed, which may limit the generalisability of the results.

BOX 6.5
EXAMPLE OF AN INTERVIEW METHOD: USE OF GUIDELINES IN PRIMARY CARE IN THE UNITED KINGDOM (Langley 1998)
The framework for interviewing 20 GPs included opening comments to elicit GPs' overall views and perceived problems; possibilities of addressing problems with an evidence-based approach; perceived need for information; information-seeking behaviour (source, storage, retrieval); guidelines as a source of information (amount, content, presentation and assimilation); evaluation of information in assisting everyday decisions; perception of guidelines as authoritative/gold standard/use in litigation; difficulties experienced (knowledge gaps and patient characteristics) and how guidelines might help.

The audiotaped interviews were transcribed and analysed using a grounded theory approach. The raw data were divided into simpler text units, which were compared in order to identify groupings of similar notions. The first analysis was on a small part of the data. This process was repeated, using more data and adjusting categories until it included as much of the data as possible.

GROUP METHODS

Group methods, like surveys and interviews, focus on perceived barriers and facilitators. The communication between group members helps to identify factors that may not be raised in individual approaches. A number of approaches have been developed, which differ with respect to their aims and the structuring of the communication between individuals in the group.

Focus group interviews

These interviews with 4-12 participants are loosely structured, focusing on 2-4 topics, and moderated by a facilitator (Morgan 1988). This method has been useful in identifying barriers and facilitators, but focus groups may be difficult to organise with busy practitioners. When feasible, focus groups can be useful as a first step to identify potentially relevant factors.

Other group methods

Brainstorming generates ideas in face-to-face interaction between participants who respond to each others' suggestions to identify new ones, without criticising any ideas. Groups of clinicians may use this technique to identify perceived barriers and facilitators for change, but practical examples related to implementation of innovations in health care

Box 6.6
FOCUS GROUP INTERVIEWS WITH GENERAL PRACTITIONERS, RHEUMATOLOGISTS AND PATIENTS ON CARE FOR RHEUMATOID ARTHRITIS (Pollemans 1996)
A written survey among general practitioners (GPs) and rheumatologists on barriers for mutual collaboration showed a difference of perspective between the two professional groups. The majority of the GPs thought that diagnosing RA is their task, while the rheumatologists thought that they should be involved. In order to understand the different views and identify opportunities for collaboration between the two professional groups, focus group interviews were performed. Seven GPs, three rheumatologists and seven RA patients were recruited. Separate meetings were organised for each of these groups. The participants received a summary of the survey findings before they attended the focus group interviews. The interviews with GPs and rheumatologists focused on their opinion about these findings. The interviews with RA patients focused on expectations on and experiences with health care. The interviews were performed by one researcher as a facilitator and another who made extensive notes. The participants were encouraged to discuss with each other, not with the researchers. The results were analysed and organised in different themes. It was concluded that general practitioners and rheumatologists have different ideas about appropriate care for RA patients.

settings were not identified. Brainstorming may be a useful element of a focus group interview.

Consensus methods, such as the nominal group technique and the Delphi procedures (Box 6.7), are used to create consensus in a group of clinicians by means of structured feedback of individuals' answers to other group members. Face-to-face interaction between participants is limited or absent. These methods can be used to determine the perceived relevance of different barriers and facilitators, so that an implementation strategy can be focused on the most relevant factors.

DIRECT OBSERVATION METHODS

Direct observation of actual events or situations may be adequate when other methods for identifying barriers and facilitators cannot be used for some reason, or when other methods need to be supplemented to

Box 6.7

A DELPHI-PROCEDURE TO IDENTIFY CONDITIONS FOR PREVENTION (Van Drenth 1998)

A Delphi procedure was used to identify organisational and administrative aspects of prevention of cardiovascular diseases in general practice. Ten general practitioners and ten practice assistants, with expertise or special interest in cardiovascular disease prevention, were invited. The members of the panel remained anonymous to each other in order to prevent direct personal communication.

A set of 29 aspects of health care was selected by the research team from the literature, from discussions with experienced primary care teams, and from visits to relevant research projects. These aspects referred to requirements for systematic case-finding and monitoring of persons at high risk of developing cardiovascular disease. In three successive questionnaires the panel members were asked to identify the most crucial requirements, taking the feasibility in day-to-day care explicitly into account. In the second and third questionnaires a selection of arguments for and against each aspect were presented and requirements were reformulated when appropriate. The members of the panel were invited to add guidelines or to formulate new requirements when considered necessary. To be selected or rejected as a practice guideline, 80% of the members of the panel had to express the same opinion on a condition. Furthermore, at least 70% of the general practitioners and at least 70% of the practice assistants had to share the opinion.

strengthen the conclusions to be drawn. However, the relevance of observed factors for actual behaviour change cannot directly be observed. Conclusions are based on comparisons between situations where change took place and those where it did not.

Participant or non-participant observation

A trained observer can be used to observe actual events and situations in a specific clinical setting. This observer may participate in the normal routines, for instance as a (simulated) patient or a care provider, or the observer may be non-participating. There are several methodological problems, such as the possibility that the observer influences the care providers and patients, if these are aware of the observations, and the possibility that the observer looses his or her independent judge-

ment, because of involvement in the situation. Observation is often time consuming and expensive, so its feasibility may be limited. Nevertheless, it can provide important information that could not be obtained with other methods.

Box 6.8
PARTICIPANT OBSERVATION OF INFORMATION TRANSFER IN PRIMARY CARE CENTRES (Diwan 1997)
The influence of situational factors on the transfer of information in hyperlipidaemia management was studied in Swedish primary care. Participant observation was done at regular intervals in two community health centres. A three-month period of participant observation at two health centres included interviews with key persons in the staff. There were repeated contacts over 18 months with 12 men at one of the health centres which participated in a screening programme for prevention of cardiovascular diseases. Data were recorded through ethnographic observation and detailed notes performed during the fieldwork period. More focused studies of 28 interactions between nurse and patients during information talks applied both observation and tape recording.

Clinician self-observation

The clinician may observe his/her own behaviour for instance by filling in a questionnaire on actual behaviour directly after a consultation. It is unclear to what extent such answers are valid; some researchers claim that this is basically a survey method. Empirical evidence on the validity of self-report is limited, but an unpublished pilot study showed acceptable correlation between self-registration, observation by a non-participating observer and scoring of audiotapes of the consultations (Spies 1998, personal communication). The method can be time consuming for the clinicians involved and therefore less feasible. The advantage of self-registration is that it can also be used to document clinicians' considerations which cannot be observed by another individual.

ANALYSIS

Barriers and facilitators for change are not always labelled as such in the data, and interpretation may be needed. In observational studies, comparisons between different situations may be needed to identify relevant factors. When barriers and facilitators have been identified, the analysis

may take different approaches: a ranking of the different factors according to their perceived importance or an analysis of the predictive value of different factors for actual behaviour or behaviour change. Furthermore, the analysis may focus on the total population of clinicians and patients studied, or on subgroups that may have different barriers and facilitators.

Conclusions

This chapter provided an overview of methods to identify barriers and facilitators for implementation of clinical guidelines or research evidence. In the absence of conclusive evidence, a combination of in-depth studies identifying barriers and facilitators with large scale studies measuring the relevance (and proportion) of these factors in specific populations of clinicians and patients seems most useful. In-depth methods often use small and qualitative data sets, for instance interviews with individuals or groups and participant observation. Large studies may use written surveys or analyse data on clinical performance. Surveys are probably most useful if they refer to concrete experiences, so case-specific questionnaires are an interesting approach.

Further research is needed on the usefulness of the different methods in this context. Most importantly, we need to know which methods help the most to design more effective implementation strategies and/or help to explain the variation of effectiveness of an implementation strategy across different units or clinicians. The current lack of evidence is an important barrier in the process of identifying barriers and facilitators for change.

Key messages

- Identification of barriers and facilitators for implementation of clinical guidelines probably helps to design more effective implementation strategies
- A range of potentially useful methods for identifying barriers and facilitators for change is available, including surveys, interviews, group interviews and direct observation.
- A combination of in-depth studies to identify barriers and facilitators and large scale studies measuring the relevance of these factors in a specific population of clinicians and patients seem most useful

Recommended literature

1. Oxman AD, Flottorp S. Chapter 8: An overview of strategies to promote implementation of evidence based health care. In: Silagy C, Haines A. *Evidence based practice in primary care.* London: BMJ Books 1998.
2. Grol R. Beliefs and evidence in changing clinical practice. *Br Med J* 1997; 315: 418-421.

References

Chinburapa V, Larson LN, Lyle Bootman J et al. Prescribing intention and the relative importance of drug attributes. A comparative study of HMO and fee-for-service physicians. *J Pharmaceut Market Management* 1987; 2: 89-105.

Cohen HR, Garwood H. Intervention to reduce telephone prescription requests. *Can Fam Phys* 1997; 43: 1952-1957.

Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. *JAMA* 1995; 274: 700-705.

Denig P, Haaijer-Ruskamp FM, Zijsling DH. How physicians choose drugs. *Soc Sci Med* 1988; 27: 1381-1386.

Denig P, Haaijer-Ruskamp FM. Descriptief besliskundig onderzoek op het gebied van voorschrijfgedrag (Descriptive decisionmaking research on prescribing behaviour). *Huisarts en Wetenschap* 1998; 41: 274-279.

Diwan VK, Sachs L, Wahlstrom R. Practice-knowledge-attitudes-practice: an explorative study of information in primary care. *Soc Sci Med* 1997; 44: 1221-1228.

Dolan JG. A survey of clinicians' opinions regarding the value of published decision analyses as sources of clinically useful information. *Med Decis Making* 1997;17: 427-430.

Grimshaw JM, Freemantle N, Wallace S, Russell I, Hurwitz, B, Watt I, Long A, Sheldon T. Developing and implementing clinical practice guidelines. *Quality in Health Care* 1995; 4: 55-64.

Grol R, Van der Weijden T, Wensing M, Te Giffel M. Effectiveness of methods and interventions to implement guidelines and change professional behaviour. Nijmegen: Centre for Quality of Care Research, 1996 (in Dutch).

Grol R. Beliefs and evidence in changing clinical practice. *Br Med J* 1997; 315: 418-421.

Langley C, Faulkner A, Watkins C, Gray S, Harvey I. Use of guidelines in primary care-practitioners' perspectives. *Family Practice* 1998; 15: 105-111.

Lomas J. Teaching old (and not so old) docs new tricks: effective ways to implement research findings. In: Dunn EV, Norton PG, Stewart M, Tudiver F, Bass MJ. *Disseminating research/changing practice.* London: Sage Publications, 1996.

Morgan DL. Focus groups as qualitative research. Beverly Hills: Sage (Sage University Paper Series on Qualitative Research Methods 160), 1988.

Newton J, Knight D, Woolhead G. General practitioners and clinical guidelines: a survey of knowledge, use and beliefs. Br J Gen Pract 1996; 46: 513-517.

Oxman AD. *No magic bullets.* London: North Thames Regional Health Authority, 1994.

Pollemans MC, Gorter S, Rethans JJ, Van der Linden S, Grol R. A matter of perspective: Final report on the project collobaration between general practitioners and rheumatologists with respect to rheumatoid arthritis. Maastricht: NHH/WOK/NVR report, 1996 (in Dutch).

Rethans JJ, Westin S, Hays R. Methods for quality assessment in general practice. *Fam Pract* 1996; 13: 468-476.

Rossignol M, Abenhaim L, Bonvalot Y, Gobeille D, Shrier I. Should the gap be filled between guidelines and actual practice for management of low back pain in primary care. The Quebec experience. *Spine* 1996; 21: 2983-2899.

Segal R, Hepler CD. Drug choice as a problem solving process. *Med Care* 1985; 23: 967-976.

Sibley JC, Sackett DL, Neufeld V et al. A randomized trial of continuing medical education. *N Engl J Med* 1982; 306: 511-515.

Van Boven K, Dijksterhuis P, Lamberts H. Defensive testing in Dutch family practice. Is the grass greener on the other side of the ocean? *J Fam Pract* 1997; 44: 468-472.

Van der Weijden T, Grol RPTM, Schouten BJ, Knottnerus JA. Barriers to working according to cholesterol guidelines. A randomized controlled trial on implementation of national guidelines in 20 general practices. *Eur J Public Health* 1998; 8: 113-118.

Van Drenth BB, Hulscher MEJL, Van der Wouden JC, Van Weel C, Grol R. Practice guidelines for cardiovascular risk screening in general practice. Nijmegen: WOK, unpublished paper, 1998.

Wensing M, Van der Weijden T, Grol R. Implementing Guidelines and Innovations in Primary Care: Which Interventions are Effective? *British Journal of General Practice* 1998; 48: 991-997.

Chapter 7

Planning, monitoring and describing interventions

Marlies Hulscher, Miranda Laurant, Michel Wensing, Richard Grol

Aims of the chapter

This chapter provides a framework and tools to plan, monitor and describe in detail the elements of a strategy for implementing guidelines in health care settings. It provides detailed examples of study plans and articles describing actual implementation studies.

Introduction

Little is known about what makes interventions to implement guidelines in daily care successful. Many potentially effective strategies to implement guidelines are available, but most studies show at best only modest improvements in performance following intervention. Most interventions were effective in some settings, but not in others. For example, feedback was effective in one study (Nattinger 1989), while in another study feedback to health professionals failed to reach positive outcomes (Szczepura 1994). Close analysis showed, however, that 'feedback' in these two studies implied different elements (Hulscher 1998). More intensive efforts to alter practice (combining various strategies) were generally more successful, but even multifaceted interventions do not always improve performance (Bero *et al.* 1998; Wensing *et al.* 1998). These results underline the importance of looking inside the 'black box' of the intervention.

To understand which elements are effective, and in what situation, it is crucial to study the processes in which successful strategies have worked or unsuccessful ones failed (Kanouse *et al.* 1995). To achieve this goal, information on the characteristics of implementation programmes is needed. Detailed information on the degree in which this programme is actually carried out is important, as defects in the implementation of the intervention may explain the lack of success. Evaluations focusing on implementation activities are useful for developing insight into the reasons for the achievement (or lack of achievement) of intervention objectives, or for forming hypotheses for later testing (Fink 1993).

Interventions to implement guidelines are, however, usually poorly described and information on the process by which an intervention works or doesn't work is scarce. As no instrument was available, the CPP framework (chapter 2 of this volume) was elaborated to enable guidelines implementation researchers to use it for planning, monitoring and describing their interventions in detail. In this chapter, we describe the development of a part of the CPP framework concerning interventions.

Methods and outline of the chapter

The literature was searched to identify relevant elements for a preliminary framework for planning, monitoring and describing an intervention. This preliminary framework was tested in three pilot studies. The first pilot study tested the feasibility and usefulness of the framework for prospectively planning an intervention and providing a detailed description of it. Based on this information, tools to monitor the elements of an intervention were developed. We piloted the CArdiovascular Risk reduction in Primary carE (CARPE) project as a test case. The third pilot study tested the usefulness and comprehensibility of the framework for a retrospective description of interventions in a series of already completed studies. Based on the results of these pilots, some alterations were made to the preliminary framework.

These steps are described in detail – and in the named order – below.

Development of a preliminary framework

To identify relevant elements for the framework, we used the CCEPP-data collection checklist, evaluation research literature, and theories on implementation and behaviour change.

The data collection checklist developed by the Cochrane Collaboration (Cochrane Effective Practice and Organisation of Care Review

Group 1998) can be used by researchers for data extraction of studies on implementation of guidelines. This checklist describes study characteristics of professional, financial, organisational, and regulatory interventions. Additional characteristics of the intervention include its content, format, source, recipient, deliverer, timing, setting, and source of funding. Classification and analysis of the intervention are done at a global level, providing little insight into the concrete elements of the implementation strategies. We used this global list of intervention characteristics as a starting point for our framework.

Interventions that are not clearly conceptualised or well-designed are difficult to evaluate. *Evaluation research literature* stresses the importance of good planning of the intervention and of appropriately set goals that are related to existing conditions (such as the perceived barriers) and to the specific setting. Selecting the target population (i.e. the group whose work the intervention is designed to improve) is the second step in the development of an intervention. The selection criteria and process need to be clearly described. Thirdly, it is relevant to look at the organisation and process of delivering the intervention, such as the schedule of activities and the accepted variation in delivering the intervention. Finally, the availability of resources (time, personnel, and materials) must be taken into account in the planning phase (King *et al.* 1987; Rossi and Freeman 1993; Stecher and Davis (1987).

From *theories on implementation and behaviour change*, underlying different approaches to implementing guidelines and changing practice, some central elements were extracted to complete the framework. *Communication theories* provide various models to explain the communication process. In most of them, the following question can be recognised: *Who* (Sender) *says What* (Message) *to Whom* (Receiver) *and How* (Medium), *When (at what moment)* (Timing) and *Where (under which circumstances)* (Noise)? Each of these elements has features that can be manipulated to increase the chance of successful communication or, in our case, implementation (McGuire 1985).

In *psychological and cognitive theories* (Ajzen 1991; Bandura 1986; Festinger 1954), social influence is considered an important determinant of behaviour change, and therefore it is a relevant element of the intervention. People learn from other people, by observing others and comparing themselves with them. They may also perceive social pressure to adapt or avoid a certain kind of behaviour.

Implementation can also be seen as a process of *diffusing innovations.* Rogers (1983) distinguished four phases in the diffusion process (dissemination, adoption, implementation and continuation). Several important determinants of the innovations (their compatibility, flexibi-

lity, reversibility, relative advantage and complexity) can be manipulated to reach successful adoption and implementation.

Based on this information, the following elements of the implementation strategies applied were considered to be potentially important in the planning, monitoring and description of interventions (Box 7.1). It is necessary to analyse each of the various types of interventions (professional, financial, organisational and regulatory interventions) separately.

BOX 7.1
PRELIMINARY FRAMEWORK TO PLAN, MONITOR AND DESCRIBE THE INTERVENTION

1. **Flexibility** means the accepted *variation* (or standardisation) in delivering the intervention (site to site/time to time).
2. **Timing** includes the time interval between delivering the intervention and clinical decision-making *(proximity)* as well as the *number* and the *duration* of intervention events and time interval(s) between these events *(frequency)*.
3. **The content** of the information consist of the *message(s)* (e.g. general or specific information on guidelines and/or performance, descriptive or graphical information), and its *comparability* (the possibility of comparing the received data on performance with those from others, or with standards).
4. **The medium** for delivering the message(s) can be for example oral, written, electronic, or a combination of these.
5. **The sender** (**deliverer**) of the message has various characteristics, including his or her *profession* (also in relation to the *clinical problem*) and perceived *authority* (credibility, attractiveness, power).
6. **The receiver of** (**or participant in**) the intervention can equally be described by *profession* (also in relation to the *clinical problem*). The number (targeted and actual) of receivers and their *motivation to participate* (voluntary, compulsory, financial support) needs description. State also if the intervention was delivered to individuals or groups, including group size, and whether the receivers can learn from each other *(social interaction)*.

Planning the intervention

Before carrying out an intervention, a detailed plan of it must be carefully documented. This framework can be used for planning the intervention, emphasising the crucial elements that need to be built in. Implementing guidelines is a step-by-step process in which different barriers can hinder progress (Grol 1997). For crucial barriers, effective measures have to be selected or developed.

The multifaceted intervention in CARPE (see page 209) included the provision of feedback, educational materials, and practical tools by trained outreach visitors during visits. The feedback intervention in CARPE is used here to illustrate how the framework can be used for describing the intervention plan in detail. The purpose of the feedback was to improve decision-making for patients with cardiovascular risk indicators or diseases. No variation was allowed across sites and times *(flexibility)*, but *the proximity of the feedback* to clinical decision-making may have varied, depending on when the general practitioner (GP) saw a relevant patient.

Eight feedback reports were provided, one for the practice organisational guideline and seven for the disease specific guidelines *(number)*, with approximately 6 months between the organisational guideline report and the first disease specific report, and one month between each disease specific feedback report *(time interval)*. The GPs received all eight reports in a total intervention period of 18 months. It is unknown how much time was devoted to reading/discussing the feedback *(duration)*.

The *content of the message* was specific information on the GP's adherence to the organisational guideline and to each disease-specific guideline *(type)*. This was given in descriptive and/or tabular formats for each guideline *(presentation)*. Individual performance was compared to the guideline as the norm (*comparability)*. The messages were in written format *(medium)*, and they were delivered by researchers from the university who also were general practitioners *(profession)*. These GPs were knowledgeable colleagues *(credibility)* from the same region *(attractiveness)* and had no control over rewards or punishments *(power)*.

The *receivers (participants)* were GPs and practical nurses *(profession)*, all members of the participating 120 practices *(number)*. Their participation was voluntary *(motivation to participate)* and each of them received individual feedback reports *(no social interaction)*.

Monitoring the implementation of intervention activities

The detailed plan of the intervention serves as a starting point for monitoring the implementation by listing the critical elements of the in-

tervention. A wide variety of techniques to gather data may be used singly or in combinations (Herman *et al.* 1987; King *et al.* 1987; Patton 1987; Rossi and Freeman 1993). The chosen instruments should be simple and feasible but sufficiently comprehensive. Monitoring the performance of the intervention activities is a major task, and often it's important to select which activities to attend to.

Depending on the available resources, information may be gathered by *surveys* (using existing records, questionnaires, or by setting up a record system), *interviews*, and by *direct observation* (on the spot observer or audiovisual material). Data can be collected during or after the intervention, from deliverers or receivers.

Using *existing records* (e.g. the pocket diary in which the outreach visitor records appointments with practices) for monitoring the intervention is an inexpensive method, providing data are of acceptable quality, easy to obtain and analyse. It's usually necessary, however, to set up a *prospective record system*. Ideally, this should be easy to maintain and useful for the staff's own purposes as well. *Interviews or self-administered questionnaires* can be structured or unstructured, with closed or open responses, depending on the goals of the data gathering. *Systematic on-site observations* can provide rich information. This method is feasible only when observation is unobtrusive.

Monitoring the intervention elements

Interventions can allow different amounts of variation across sites or across time. The more variation allowed, the more attention must be paid to monitoring the activities. *Flexibility* can apply to all elements in the framework. Examples of monitoring from the CARPE-project can be found under all of the following subheadings.

Timing. *For proximity to the decision-making moment*, self-reports by receivers is probably the most practical method of monitoring. Several strategies (record keeping, self-reports, and observation) can be used to monitor the *number* of events; the *time interval* between them; and the *duration* of intervention activities. (Box 7.2 opposite).

Content and medium. To monitor the content of the intervention activity, *written* information can be gathered by asking the sender to give a copy of the materials used. The characteristics of the information can then be extracted and scored on a structured checklist. Information of *oral presentations* can be monitored by an observer who records information on a checklist on the spot, or afterwards from an audio or video recording. Information can also be gathered from senders and/or recei-

Box 7.2
Prospective, structured data collection on the implementation in CARPE

PLAN: General practices were to receive ±16 tailored (flexible) outreach visits
MONITORING: Information was gathered from the providers during the intervention. A simple, coded registration form was developed and filled in by the outreach visitor after each visit to a practice, including the date and the duration of a visit. This allows for counting the actual number of visits and the time intervals between them.

vers using a structured interview or a self-administered questionnaire. The *medium* by which information is provided is easy to record together with the content. (Box 7.3.)

Box 7.3
Using checklists during intervention in CARPE

PLAN: During the first outreach visit, standardised general information about the project, the guidelines and the intervention were to be given.
MONITORING: The outreach visitors were provided with a checklist of the items that had to be mentioned during the first visit. They were trained to use this checklist and carefully instructed to utilise it for each practice visited. During or after the visit, they marked the items they actually had covered. The checklists were collected and information from them extracted for the study report.

Sender (Deliverer). Establishing the background of the sender *(profession and clinical problem)* and his/her *authority* is important especially from the point of view of the receiver. Disagreement on the sender's authority between sender and receiver may be an important factor in explaining the (un)successfulness of the intervention (McGuire 1985; Thomson *et al.* 1997). Information on both topics can best be collected by means of self-reports or observation.

Receiver (Participant). To record whether the intervention was targeted at an individual or a group *(social interaction)*, the *profession* of the participants, and the *actual group size*, existing records (minutes of meetings, certificates of attendance, and sign-in/sign-out sheets) can be used. Alternatively, the participants can be asked about their participation. Interviews, questionnaires or diaries can be used to record participation in the planned activities. To test whether an intervention was received by participants, information can also be generated indirectly by testing what the participants have learned. Setting up a record system,

Box 7.4
Registering participation in CARPE interventions

PLAN: General practices were to participate in outreach visits.
MONITORING: The outreach visitor filled in a simple, coded registration form after each visit to a practice, registering the participants in the meeting (name and status). This also provided data on whether the intervention was aimed at individuals or groups and the actual group size.

or having an observer at meetings record the participants, are also options. For monitoring the *motivation to participate*, information from senders and/or receivers by self-reports is most practical. (Box 7.4.)

Describing the intervention

The framework provides an accurate account of the intervention as actually performed. This is essential to interpreting the results of the intervention: which aspects were effective, does a dose response relation exist, etc. A description of the intervention is important for determining its critical elements and as a basis for others who want to replicate the intervention in other settings. Finally, using the framework retrospectively to describe an intervention as actually performed facilitates comparisons between implementation studies. We tested the applicability of the preliminary framework for retrospective purposes.

First, we took a convenience sample of twenty-nine studies (26 authors) which had used different implementation strategies: educational materials, conferences or meetings, small group quality improvement, outreach visits, feedback, reminders, organisational interventions and financial interventions. We structured the information provided in reports of these studies using the framework. Subsequently, we sent this

overview to one of the authors and asked them to complete the framework. The authors could correct us and/or give additional information on all elements of the intervention. Two reminders were sent to those authors who had not responded. Information from nineteen of the contacted authors (73%), covering 21 studies (72%), could be used for further analysis (Axt-Adam *et al.* 1993; Browner *et al.* 1994; Buntinx *et al.* 1993; Cecchini *et al.* 1989; Cockburn *et al.* 1992; Coulter and Bradlow 1993; Cowan *et al.* 1992; Dietrich *et al.* 1992; Emslie *et al.* 1993; Gilio *et al.* 1993; Jones *et al.* 1993; Lassen and Kristensen 1992; Mandelblatt *et al.* 1993; Oakeshott *et al.* 1994; Szczepura *et al.* 1994; Steffensen *et al.* 1997; Stokx *et al.* 1993; Soumerai *et al.* 1993; Vedsted *et al.* 1997; Winkens *et al.*1995; Zaat *et al.*1991).

BOX 7.5
DESCRIBING FLEXIBILITY (Axt-Adam 1993)

ARTICLE: The laboratory in Delft regularly arranges educational meetings for GPs, discussing test ordering in specific diseases (e.g. thyroid or renal function) to influence the test ordering behaviour. Written reports of these "round-table conferences" are sent to all GPs in the region, regardless of their attendance.
ADDITIONAL INFORMATION: The monthly ***educational meetings*** were not very standardised. Every educational meeting was quite unique, addressing a specific topic. Furthermore, the content of the educational meetings would depend on the participants, and on the expert who co-chaired it.

Flexibility. Most articles (81%) supplied no information on the allowed site-to-site or time-to-time variation. However, all researchers were able to tell us whether the intervention had been standardised or flexible, and to what extent variation was allowed (Box 7.5).

Timing. *Proximity to clinical decision-making*: Some information on the time interval from intervention to clinical decision-making was found in seven implementation studies using feedback, reminders or organisational interventions. Eleven authors could provide adequate additional information on the timing (e.g. Box 7.6).

BOX 7.6
DESCRIBING TIMING (Dietrich *et al.* 1992)

ARTICLE: One part of an intervention to improve prevention and early detection of cancer was ***outreach visits*** by a project facilitator, supporting the establishment of routines for providing needed services. Facilitators visited each practice three times over three months, and provided additional assistance as needed.
ADDITIONAL INFORMATION: The outreach visits were delivered over a period of weeks-months and might begin to have impact after the very first visit. Its full impact would have been achieved 2-3 months later at the earliest. By that time, the physician would have a flow sheet ***reminder***, possibly provided by a medical assistant, pointing out that a service was indicated.

BOX 7.7
DESCRIBING TIMING (Browner *et al.* 1994)

ARTICLE: To improve the management of high serum cholesterol levels GPs were invited to participate in a ***Continuing Medical Education (CME)*** meeting. The standard CME group was offered one free 3-hour seminar. The intensive CME group was offered, in addition, two follow-up seminars and free office materials. They also received two office visits and encouraging phone calls between the seminars. After the first seminar, a staff member visited the office to explain the use of the educational materials. One month later, physicians attended a two-hour follow-up seminar, and a month after that they received a second office visit. Two to three months later, GPs were invited to attend the final seminar.
ADDITIONAL INFORMATION: Intensive CME: *Number* of phone calls varied from 2 to 5. *Time intervals* between first seminar and first office visit as well as between the encouraging phone calls varied from 1 to 3 months. The *duration* of the office visits varied from 10 to 30 minutes and the phone calls lasted from 1 to 5 minutes.

The number of intervention events, the time interval between them and the duration of each event were found in about 70% of the papers, if not always complete. Nineteen researchers provided additional information (Box 7.7). The duration of some intervention events (for example reading educational materials or feedback reports) can not be described, unless the exact time investment of care providers is monitored.

Content. *Message:* Most papers (95%) described the *content of the information* in the guideline or change proposal given. Seven papers gave a clear classification of the type of information (Box 7.8), and three studies described how the information was presented (e.g. in graphical format). Four researchers could provide additional data on the type of information, but none on its presentation. A reason for this may be that the open-ended question to the authors was not specific enough.
Comparability of the information: Only four articles (19%) described whether the information could be compared with other persons or with guidelines. Seventeen researchers could provide us with extra information (Box 7.9).

Box 7.8
Describing information content (Steffensen *et al.* 1997)

ARTICLE: A guideline on anticoagulant therapy to prevent stroke in patients with atrial fibrillation was implemented in primary care. The local guideline was mailed as a two-page 'Newsletter', and later a reminder in a local periodical was published (**dissemination of educational materials**). The doctors also received information about which patients anticoagulant therapy was recommended for, the intensity of anticoa-gulation, individual risk estimates, and suggestions for investigations. In addition, the doctors received practical suggestions on how to initiate oral anticoagulation, and information about how to prepare and mail blood samples to the laboratory for monitoring the anticoagulant therapy.
ADDITIONAL INFORMATION: None

Box 7.9
Describing comparability of data (Stokx *et al.* 1993)

ARTICLE: GPs could enter an integrated ***CME programme***, consisting of ***educational meetings, small group quality improvement, and feedback***. The goal was to improve the quality of care given by GPs. They discussed referrals to medical specialists, prescribing medication, referrals of patients with visual problems, and the diagnosis, treatment and prevention of asthma.
ADDITIONAL INFORMATION: The general practitioners also received information (feedback) on individual and group performance.

Medium. The various media used in implementation studies are seldom described in the articles explicitly. However, in 19 studies there was sufficient implicit information allowing classification of the medium. Nine researchers provided us with additional information (Box 7.10). An explicit description of the media is not always necessary. Only if the relationship between the medium and implementation strategy is unclear, it is relevant to explicitly describe the medium (see also paragraph 7.7).

Box 7.10
Describing the media in an intervention (Browner *et al.* 1994)
(see also Box 7.7 for a further description of this article)

ARTICLE: To improve the management of high serum cholesterol levels GPs received free educational materials (e.g. "Eater's Choice" (book); "A change of Heart" (patient manual); pocket-size laminated cards; postcard reminders; "Current Medical Diagnosis and Treatment" (book), and a videotape.
ADDITIONAL INFORMATION: The videotape was not for the GPs, but for the patients.

Sender (Deliverer). *Profession*: In fifteen articles, some characteristics of the sender (deliverer) were described. Twelve researchers gave additional information on the profession of the sender, so that this information was available for almost all studies.

Authority (credibility, attractiveness, power): Most articles were not very clear about these characteristics of the sender. Only one publication described the (perceived) authority of the sender. From seven articles we could derive that the credibility of the sender was 'high', and from two that the sender was perceived as attractive. Nine researchers gave additional information on the authority of the sender; one stated that the sender was chosen for her high credibility. When asked, all provided extra information about the credibility and three about the attractiveness of the sender (example Box 7.11). None described the power of the sender. A possible reason for this lack of precise information on authority may be that the open-ended questions were not specific enough.

Box 7.11
Describing the authority of the sender (Szczepura *et al.* 1994)

STUDY (PAPER): ***Feedback*** on screening, immunisation etc. at the practice level was used in an intervention; some of it was delivered by medical facilitators. These facilitators were nominated by the local medical committee, and included the committee secretary, a committee member, and two academic general practitioners.
ADDITIONAL INFORMATION: Practices perceived facilitators to be authority figures, but the facilitators did not instruct the practice to change in any particular way. Researchers delivered the results (feedback reports), and accompanied the facilitators on their visit.

Receiver (Participant). *Profession*: All publications described to whom the intervention was delivered, e.g. hospital doctors, practice staff. Seven researchers provided us with additional information about the professions of the participants, e.g. the professions of the different practice staff members.

Targeted number: Eighteen studies mentioned how many care providers or patients received the intervention. None of them related the number of participants to the targeted number (sample). The researchers who did not mention numbers in the article could provide the missing information.

Social interaction: Fourteen articles mentioned whether the intervention was delivered to groups or individuals. Eight studies described interventions delivered to groups, and in four of them the group size was stated. Eleven researchers gave additional information on this aspect.

Motivation to participate: Two articles reported the motivation to participate; in both studies, the participants received free educational materials, and in one of these a compensation was offered. Eighteen researchers could provide us with additional information (e.g. Box 7.12); they reported whether participation was voluntary (n = 10) or compulsory (n = 2), and if participants received free materials (n = 4) or a financial compensation to participate in the study (n = 2).

BOX 7.12
INFORMATION ON PARTICIPANTS (Cowan *et al.* 1992)

STUDY: A fact sheet ***reminder*** provided information about seven periodic health examination actions (mammography, cervical smear, faecal occult blood, cholesterol history, and vaccinations for tetanus, pneumonia, and influenza). Participants were 16 first-year medical residents, providing care for a group of patients, and serving as their primary care physicians.
ADDITIONAL INFORMATION: *Social interaction:* Individual. *Motivation to participate:* Participation in the study (intervention/control) was compulsory, but compliance with preventive interventions voluntary.

CONCLUSIONS

We have generated, tested and evaluated a framework to be used by implementation researchers as an instrument to plan, monitor and describe in detail the elements of an intervention programme for implementing guidelines in health care settings. This is incorporated as a part of the interventions sections in the CPP framework for guideline implementation studies (chapter 2).

The framework we developed was useful for interpreting intervention studies in which different strategies were applied. Further research is necessary, however, to identify relevant elements for describing financial interventions. When using the framework to monitor interventions, it may be important because of resource restraints to choose which elements of the intervention activities to pay attention to. Some

elements (medium, authority etc.) are best monitored by one simple, structured question to senders or receivers. For example, the media used to distribute educational materials can be classified to (i) paper (e.g. leaflets, books); (ii) audio/videotapes; or (iii) computer software (e.g. CD-ROMs). Finally, our description of some elements in the framework was not specific enough (e.g. the content of the information or the authority of the sender), as detailed data on these elements were often lacking even after asking the authors to provide it. Based on these results we adjusted the framework (see Box 7.13, page 148).

Our approach had some limitations. The literature review to identify important elements of an intervention was unsystematic and may not have identified all possible elements. Planning and monitoring the intervention were tested using only one case (the CARPE project); the feasibility of using the framework for different purposes needs to be demonstrated later. The articles in the retrospective pilot study were a convenience sample of studies that were easily accessible; to cover all strategies, we made sure that the selection included professional, financial, organisational and regulatory interventions. The provision of additional explanatory information from the authors was through written communication instead of discussion, leaving some points open.

Although the framework still needs fine-tuning (which the Centre for Quality of Care Research tends to do in the coming years), the present stage can well be used to plan, monitor and describe interventions. Detailed information on the elements of an intervention makes it possible to compare studies that implement clinical guidelines. This comparison allows insight into the reasons for the achievement of intervention objectives. Detailed information on the intervention is also important for economic evaluations, allowing estimates of the cost (in terms of time and/or money) of interventions. For example, knowing the necessary number of outreach visits, their duration and the profession of the outreach visitor one can calculate the costs of such interventions.

Understanding the determinants for success or failure in the achievement of intervention objectives, and information necessary for the estimation of intervention costs are both crucial, when we want to design and implement successful and cost effective interventions facilitating the use of guidelines in health care.

Box 7.13
Framework for planning, monitoring and describing an intervention

Relevant elements of the intervention	Questions to elicit the information[a]
1) Implementation strategy or strategies used[b] • *Professional interventions* • *Financial interventions* • *Organisational interventions* • *Regulatory interventions*	***1a)*** Could you describe which implementation strategies you are to use in your implementation study? ***1b)*** Could you describe your strategies in some detail? Which activities have you planned or performed (see questions 2 to 7)?
The following elements must be described separately for each of the implementation strategies.	
2) Flexibility • *Variation*	***2)*** Could you describe if it is allowed to alter the intervention across different practices/sites or across time? Describe this variation, if possible.
3) Timing • *Proximity to decision-making* • *Number* • *Frequency (time interval)* • *Duration*	***3a)*** Could you describe the timing of the intervention, i.e. the interval from the intervention event to clinical decision-making (in a consultation)? The intervention is *prospectively linked* to the decision, e.g. when the computer programme reminds the GP what drug should be prescribed at the moment the GP wants to write a prescription; *retrospectively linked* when directly after the consultation a facilitator tells the GP what he has observed in the consultation and gives practical advice for improving the performance; and *unlinked* when the GP receives a feedback report at a later date. ***3b)*** Could you describe the number of similar intervention events (e.g. two feedback reports, four educational meetings)? ***3c)*** Could you describe the intervals between similar intervention events e.g. 6 months between the first and second educational meeting, a feedback report every month)? ***3d)*** Could you describe the duration of each similar intervention event (e.g. two-hour educational meeting, etc.)?

a) Depending on the purpose the framework is used for, the questions need to be formulated in the present tense (planning or monitoring the intervention), or past tense (describing the intervention).

b) See chapter 2 for the entire CPP framework for guideline implementation studies.

4) Content • *Type of information:* general/specific information innovation, general/specific information performance • *Presentation of information:* descriptive, illustrations, graphical, tabular • *Comparability of the information:* with others, with standards	***4a)*** Could you describe the type of information delivered in the intervention (e.g. general advice on individual performance, specific information on guideline recommendations on laboratory samples, patient specific advice)? ***4b)*** Could you describe how the information is presented in the intervention (e.g. descriptive, illustrations, graphical)? ***4c)*** Could you describe the possibility of comparing one's behaviour or performance with others or to common criteria (e.g. comparable with colleagues in the same region, with opinion leaders, with national evidence based guidelines)?
5) Medium • *Oral, written, combination*	***5)*** Could you describe how the intervention is delivered (e.g. personal discussion, telephone, paper, computer)?
6) Sender • *Profession (patients: clinical problem)* • *Authority* *1. credibility* *2. attractiveness* *3. power*	***6a)*** Could you describe who or what delivers the intervention (e.g. local expert, researcher, colleague)? State also the profession of the sender (for patients, their clinical problem). ***6b)*** Could you describe the perceived authority (i.e. (1) credibility, (2) attractiveness and (3) power) of the sender(s)? For example, high credibility due to the sender's knowledge, high attractiveness because the sender is familiar to the receiver, low power because the sender gives no rewards or punishments.
7) Receiver (participants) • *Social interaction:* individual, group (size) • *Profession (patients: clinical problem)* • *Number (targeted/actual)* • *Motivation for participation:* voluntary or compulsory participation, financial support	***7a)*** Could you describe if the intervention will be delivered to individuals or to a group? State the size of the group or groups. ***7b)*** Could you describe the target population of the intervention? State the profession or clinical problem of the receiver. E.g. clinicians in primary care, nurses, dentists, etc. ***7c)*** Could you describe the targeted (actual) number of participants in the intervention? ***7d)*** Could you describe the participant's motivation to join in the intervention?

KEY MESSAGES

- This framework can be used for planning, monitoring and describing an intervention in detail, although it needs new insights for future fine-tuning.
- It is necessary to describe the specific elements of interventions in detail, in order to determine which elements are crucial in the success or failure of the intervention.
- Various methods for monitoring the implementation of intervention activities are available.
- Understanding the reasons for the achievement of intervention objectives helps in designing successful interventions for guideline implementation.

REFERENCES

Ajzen I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes,* 1991; 50: 179-211.

Axt-Adams P, Wouden, JC van der, Hoek H *et al*. Het effect van nascholing op het aanvragen van laboratoriumdiagnostiek door huisartsen. *Huisarts en Wetenschap,* 1993; 36(13): 451-454.

Bandura A. *Social foundations of thought and action: a social cognitive theory.* New York, Prentice-Hall, 1986.

Bero L, Freemantle N, Grilli R et al. on behalf of the Cochrane Collaboration on Effective Professional Practice. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote implementation of research findings by health care professionals. *British Medical Jorunal* 1998; 317: 465-468.

Browner WS, Baron RB, Solkowitz S et al. Physician Management of Hypercholesterolemia. A Randomized Trial of Continuing Medical education. *WJM* 1994; 161(6): 572-578.

Buntinx F, Knottnerus JA, Crebolder, HFJM et al. Does feedback improve the quality of cervical smears? A randomized controlled trial. *British Journal of General Practice* 1993; 43: 194-198.

Cecchini S, Grazzini G, Bartoli D et al. An attempt to increase compliance to cervical cancer screening through general practitioners. *Tumor* 1989; 75: 615-618

Cochrane Effective Practice and Organisation of Care Review Group (EPOC): *The Data Collection Checklist*. 1998

Cockburn J, Ruth D, Silagy C *et al.* Randmised trial of three approaches for marketing smoking cessation programmes to Australian general practitioners. *British Medical Journal* 1992; 304: 691-694.

Coulter A, Bradlow J. Effect of fundholding and indicative prescribing schemes on general practitioners' prescribing costs. *British Medical Journal* 1993; 307: 1186-1189.

Cowan JA, Heckerling PS & Parker JB. Effect of a Fact Sheet Reminder on Performance of the Periodic Health Examination: A Randomized Controlled Trial. *American Journal of Preventive Medicine* 1992; 8(2): 104-109.

Dietrich AJ, O'Conner GT, Keller A et al. Cancer: improving early detection and prevention. A community practice randomised trial. *British Medical Journal* 1992; 304: 687-691.

Emslie C, Grimshaw J & Templeton A et al. Do clinical guidelines improve general practice management and referral of infertile couples. *British Medical Journal* 1993; 306: 1728-1731.

Festinger L. A theory of social comparison processes. *Human Relations* 1954; 7: 117.140.

Fink A. *Evaluation fundamentals guiding health programs, research, and policy.* Newbury Park CA: Sage Publications Inc, 1993.

Gilio G, Buntinx F, Kezel O de et al. The Influence of a Desk-top Analyser on the Number of Laboratory Tests used in daily General Practice. A Randomized Controlled Trial. *Family Practice* 1993; 10(2): 118-123.

Grol, R. Implementing guidelines in general practice care. *Quality in Health Care* 1992; 1: 184-191.

Herman JL, Morris LL, Fitz-Gibbon CT. *Evaluator's Handbook.* Program Evaluation Kit. Herman JL (ed.). Newbury Park CA: Sage Publications Inc, 1987.

Hulscher MEJL (1998). Implementing prevention in primary care: a review. In: *Implementing prevention in general practice: a study on cardiovascular disease* (Thesis). Nijmegen: 25-70.

Jones RH, Lydeard S, Dunleavey J. Problems with implementing guidelines: a randomised controlled trial of consensus management of dyspepsia. *Quality in Health Care* 1993; 2: 217-221.

Kanouse D, Kallich J, Kahan J. Dissemination of effectiveness and outcomes research. *Health Policy* 1995; 34: 167-192.

King JA, Morris LL, Fitz-Gibbon CT. *How to assess program implementation.* Program Evaluation Kit. Herman JL (ed). Newbury Park CA: Sage Publications Inc, 1987.

Lassen LCHR, Kristensen FB. Peer Comparison Feedback to Achieve Rational and Economical Drug Therapy in General Practice: A Controlled Intervention Study. *Scand. J. Prim. Health Care* 1992; 10:76-80.

Mandelblatt J, Traxler M, Lakin P et al. A Nurse Practitioner Intervention to Increase Breast and Cervical Cancer Screening for Poor, Elderly Black Women. *Journal of General Internal Medicine* 1993; 8: 173-178.

McGuire WJ. Attitudes and attitude change. Pp. 258-294 in: Lindsay G, Aronson E. *The handbook of social psychology* (3rd ed.). New York: Random House, 1985.

Nattinger AB, Panzer RJ, Janus J. Improving the utilization of screening mammography in primary care practices. *Arch Intern Med* 1998; 149: 2087-2092.

Oakeshott P, Kerry SM, Williams JE. Randomized controlled trial of the effect of the Royal College or Radiologists' guidelines on general practitioners referrals for radiographic examination. *British Journal of General Practice* 1994; 44: 197-200.

Patton MQ. *How to use qualitative methods in evaluation.* Program Evaluation Kit. Herman JL (ed.). Newbury Park CA: Sage Publications Inc, 1987.

Rogers E. *Diffusion of innovations.* New York: Free Press, 1983.

Rossi PH, Freeman HE. *Evaluation: a systematic approach.* Newbury Park CA: Sage Publications Inc, 1993.

Szczepura A, Wilmot J, Davies C et al. Effectiveness and cost of different strategies for information feedback in general practice. *British Journal of General Practice* 1994; 43: 19-24.

Stecher BM, Davis WA. *How to focus an evaluation.* Program Evaluation Kit. Herman JL (ed.). Newbury Park CA: Sage Publications Inc, 1987.

Steffensen FH, Sørensen HT, Olesen F. Impact of local evidence-based clinical guidelines – a Danish intervention study. *Family Practice* 1997; 14(3): 209-215.

Stokx LJ, Gloerich ABM, Hoenen JAHJ et al. Kostenbesparing door kwaliteitsbevordering. De effecten van geïntegreerde nascholing voor huisartsen. *Huisarts en Wetenschap* 1993; 36(13): 440-444.

Soumerai SB, Salem-Schatz S, Avorn J et al. A Controlled Trial of Educational Outreach to Improve Blood Transfusion Practice. *JAMA* 1993; 270(8): 961-966.

Thomson MA, Oxman AD, Haynes RB, Davis DA, Freemantle N, Harvey EL (1997). Local opinion leaders to improve health professional practice and health care outcomes (Cochrane Review). In: *Cochrane Library*, Issue 4, 1998: Oxford: Update Software.

Vedsted P, Nielsen NJ, Olesen F. Does a computerized price comparison module reduce prescribing costs in general practice? *Family Practice* 1997; 14(3):199-203.

Wensing M, Weijden T van der, Grol R. Implementing guidelines and innovations in general practice: which interventions are effective? *British Journal of General Practice* 1998; 48: 991-997.

Winkens RAG, Pop P, Bugter-Maessen AMA et al. Randomised controlled trial of routine indidivual feedback to improve rationality and reduce numbers of test requests. *Lancet* 1995; 345: 498-502.

Zaat JOM, Eijk JthM van & Bonte HA. Mag het ook een testje minder? De invloed van een beperking van het aanvraagformulier voor laboratoriumonderzoek. Huisarts en Wetenschap 1991; 34(2): 72-77.

Chapter 8

Attitudes towards guidelines and a scale for measuring them

Marko Elovainio, Martin Eccles, Marjukka Mäkelä

Aims of the chapter

The aim of this chapter is to discuss why attitudes of health care professionals towards clinical guidelines need to be considered in guidelines implementation studies, and to describe the development and validation of the Attitudes towards Guidelines Scale (AGS) for measuring these attitudes.

Introduction

Considerable interest exists in developing and implementing clinical practice guidelines (Grimshaw *et al.* 1993). Although guidelines are generally welcomed, clinicians often do not change their behaviour in line with the recommendations (Grol 1990). Potential reasons for these results are the characteristics of guidelines (Hellbruck 1997), and the attitudes of health care providers towards guidelines and implementation strategies (Mansfield 1995).

Most research and reviews have concentrated on analysing factors concerning the perception of guidelines and characteristics of the implementation process. According to these results, reasons for the resistance to guidelines are various, but most often they are associated with the perceived characteristics of guidelines, such as unreliability or impracticality of guidelines. Professionals may also fear that guideline recommendations reduce their autonomy at work (Grimshaw *et al.* 1993;

Kanouse *et al*. 1988; Rappolt 1997), or that clinical guidelines are being created only to reduce health care costs. When implementing clinical guidelines, it is essential to understand which expectations and fears health care professionals may have toward these guidelines. Knowing such attitudes makes it possible to consider them in planning an effective implementation strategy. Although it may be difficult (or unnecessary) to affect attitudes as such, implementation strategies can be designed to overcome practical problems expressed through attitudes and to help strengthen the team approach often needed in guideline implementation. For these reasons, a scale for measuring attitudes toward guidelines was needed for European use.

Aims and methods

Within the Concerted Action, we wanted to develop and test a survey instrument with acceptable psychometric properties which would cover the most essential perceived barriers and carriers concerning the implementation of guidelines. First, we wanted to cover the area of perceived barriers to and facilitators of guideline implementation as completely as possible. We wanted the instrument to measure individual differences in potential users, their work characteristics, and variations in organisational climate.

These aims were slightly modified for two reasons. Results of previous research suggest that the perception of guidelines is the most central factor in health professionals' attitudes towards guidelines. Within Changing Professional Practice (CPP), serious doubts were also expressed about the appropriateness of an extensive survey instrument for this kind of evaluation. We decided to develop a short scale covering the main factors behind attitudes toward guidelines.

The aim was to measure attitudes about the following aspects of guidelines: their usefulness, practicality, and reliability; the availability of guidelines (Varonen et al. 1997), and individual and organisational competence to use guidelines. For each of these aspects, several items were used as either positive or negative (reverse) statements. The actual use of guidelines was indicated by one item, 'I use guidelines regularly at my work', answered using the same scale. The questions were presented as part of a more extensive survey instrument used in the partner project, covering personality variables, job characteristics, team climate and strain in primary care.

Instruments

Using previously generated lists of questions, we compiled as complete

an instrument as possible. Perceived factors or attitudes, as we call them in this chapter, affecting the use of guidelines were assessed with a 39-item scale, based on scales developed in RAND and Stakes (National Research and Development Centre for Welfare and Health). Some items were based on the questions used in research by Mansfield (1995). Responses to the items were given on a seven-point scale ranging from strongly disagree (1) to strongly agree (7).

Samples

The study population was from one CPP partner project (Mäkelä *et al.* 1998), aimed at improving the quality of care and the work environment in primary health care settings. The samples consisted of all workers in 12 health centres (HCs) around Finland sampled in two batches, six HCs starting the study in 1996 and six in 1998. The health centres varied in size and represented different areas in Finland.

Sample 1:

A postal questionnaire sent to six HCs in April 1996 was completed and returned by 748 respondents (66% of the staff). Of the respondents 95% were women, 16% were physicians or dentists, 31% registered nurses and 27% practical nurses. For the purpose of developing the final questionnaire, only responses from physicians, registered nurses and practical nurses were used (n=499). The mean age was 41.6 years (SD 9.1 years). The average duration of the present employment was 10.3 ± 7.5 years. All of these figures are comparable with those of previous random samples of Finnish healthcare providers.

Sample 2:

A postal questionnaire to the other six HCs was sent in April 1998, and was completed and returned by 688 respondents (63%). Of the respondents 9% were physicians or dentists, 25% registered nurses and 18% practical nurses. Again only the responses of physicians and of registered and practical nurses (n=358) were used for this study. Their mean age was 43.2 ± 8.7 years and the mean duration of employment 12.0 ± 7.9 years; again, all these figures are comparable to previous samples.

Statistical Analysis

The development of a short instrument for measuring attitudes towards guidelines comprised of four sets of analyses which are described in more detail in Annex 1 at the end of this chapter. We first did an explora-

tory analysis of the structure of the whole scale and dropped items which did not fit any of the tested structures. Next we performed a confirmatory factor analysis (LISREL 8.30) on the remaining items to explicate the final structure and to further reduce the number of items. The internal consistency of the questionnaire and the predictive validity of the final scale were then tested.

Results

Exploratory factor structure of the Attitudes towards Guidelines Scale (AGS)

A six-factor solution was the best fit in the exploratory factor analysis when the factors were not allowed to correlate. This solution partially supported the hypothesised structure of the scale. Some items did not load most highly on their theoretically predicted factor. The six-factor model explained 55% of the total variance in the first phase (with 39 items) and 63% in the second phase (with 27 items).

Based on that structure the following six subscales were developed and analysed:

1) usefulness of guidelines, 2) reliability of guidelines, 3) lack of individual competence to use guidelines, 4) lack of organisational competence to use guidelines, 5) impracticality of guidelines, and 6) availability of guidelines. The subscales with their items and reliability (Cronbach's alpha) are given in Annex 2 of this chapter.

Factor reduction of the Attitudes towards Guidelines Scale (AGS)

From the full model of six subscales, including 27 of the original items, items were excluded one by one following modification indices provided by LISREL. In the first acceptable measurement model the number of items was reduced to 14 and one factor was split into two, producing seven factors. Most indices we used supported the conclusion that this model provided a reasonably good fit to the data (see Annex 1 of this chapter). The LISREL programme itself suggests changes in the model that may improve the fit. However, none of the suggested changes significantly improved the fit.

The final structure was independently tested in sample 2, and the model provided a reasonable good fit to this other data as well. Based on the final structure the seven subscales (two items each) were formed and analysed. The scoring of subscales 4, 5 and 6 were reversed, and a composite general score was computed. The subscales and their items as well as the model with its factor loadings are presented in Table 8.1. The factor loadings essentially indicate that most of the observed items

Table 8.1. Attitudes Towards Guidelines Scale: *Subscales,* their reliability (Cronbach's alpha, from Sample 1 / Sample 2), and their items with factor loadings (from Sample 1 / Sample 2).

General attitude towards guidelines (0.88 / 0.79)
Guidelines are useful as educational tools. (0.90/0.32)
Guidelines are a convenient source of advice. (0.87/0.54)
Usefulness (0.63 / 0.72)
Guidelines can facilitate communication with patients and families. (0.72/0.58)
Guidelines can improve the quality of health care. (0.74/0.48)
Reliability of guidelines (0.83 / 0.79)
Guidelines are based on scientific evidence. (0.97/0.74)
Guidelines are made by experts. (0.79/0.86)
Lack of individual or team competence (0.50 / 0.42)
My occupational competence is insufficient for adopting the latest guidelines. (0.51/0.69)
Most of our team members have disapproving attitudes about guidelines. (0.78/0.81)
Lack of organisational competence (0.63 / 0.45)
Guidelines are not valued in our organisation. (0.69/0.83)
To implement guidelines is too expensive for us. (0.68/0.79)
Impracticality (0.68 / 0.69)
Guidelines challenge the autonomy of care providers. (0.91/0.73)
Guidelines oversimplify medical practice. (0.53/0.71)
Availability (0.57 / 0.66)
Guidelines are difficult to find if needed. (Reversed) (0.84/0.72)
I have not seen any guidelines in our health care unit. (Reversed) (0.34/0.68)

measured the constructs behind them at least acceptably. The most serious exception was the last item in Table 8.1 (local availability), with the lowest loading.

Relationships between subscales and predictive validity of AGS

Annex 3 of this chapter presents the basic statistics on the measures used in this study. The figures show that the attitudes of the HC staff towards guidelines in general were rather positive, and that all of the AGS subscales were related to each other. The respondents' views of the reliability and usefulness of clinical guidelines were also quite positive. On the other hand, they did not think very highly about the availability or practicality of the guidelines or their competence for using them. There were no statistically significant differences between occupational groups, although physicians and dentists more often reported that their individual competence to use guidelines was appropriate (Annex 4 of this chapter).

To test the predictive validity of the scale, the relationship between AGS subscales and the reported use of guidelines was analysed. The regression model, adjusted by age, occupation and tenure, showed that

perceptions about the availability of guidelines and individual/team abilities to use guidelines were significant predictors of the reported use of guidelines in both samples. In sample 2, the perceived reliability of guidelines also was a significant predictor of reported guideline use (Annex 5 of this chapter).

Discussion

In this study we developed and tested the structure and psychometric properties of the Attitudes towards Guidelines Scale (AGS). This scale was based on previous studies done by Stakes and RAND and also on theoretical discussions (Eccles *et al.* 1996; Eddy 1990; Feder *et al.* 1999; Grimshaw *et al*. 1993 and 1998; Hellbruck 1997; Kanouse *et al*. 1988; Rappolt 1997). According to our results, the important factors behind a general positive or negative attitude towards guidelines are the usefulness, reliability, practicality and availability of the guidelines. In addition to these aspects, the individual/team and organisational competence to use guidelines and to follow the procedures they recommend seemed to be vital. All of these aspects could be measured with the instrument (AGS) reported here.

From the exploratory factor analysis it can be concluded that the six-factor solution of the AGS constituted an acceptable fit to the data in both samples. Exploratory factor analysis lets items load on factors other than their designated one. When the scales are subjected to further analyses, only the items with higher loadings are used. In confirmatory factor analysis items are not allowed to load anywhere but on their corresponding factors or latent variables, and the overall fit of the model to the data is judged.

The more fixed parameters and the more items that are used, the more possibilities for correlated residuals. This is stringent but congruent with the theoretical predictions of the seven-factor model. The structural equations of the items kept in the model were quite strong for both samples and their residuals were relatively small. All of the equations were statistically significant. The modification indices did not give any reasonable new solutions. Although the fit indices were not as good as might be desired, the solution was acceptable for the data in both samples.

As a statistical method LISREL offers some preferable properties compared to exploratory factor analysis. LISREL gives the opportunity to test more explicitly the factor structure as a whole, and provides a lot of information for evaluating the adequacy of the assumed model. Standard errors and correlations of the parameter estimates, measures of variation accounted for, overall goodness-of-fit measures, analysis of resi-

duals and model modification indices are all offered as an output file. When the observed variables are allowed only to load on the corresponding latent variable, the statistical characteristics of the variables become more meaningful.

Our analysis offered some information about the predictive validity of the developed scales. The relationships between the AGS subscales and the self-reported use of clinical guidelines were strong. The basic statistical analysis showed that Finnish health care professionals perceived clinical guidelines as useful and reliable, but not as very practical or easily available in their organisations. They held quite positive general attitudes towards guidelines, but thought that they themselves, or their teams or organisations, did not always have the resources or competence required to follow the procedures recommended in guidelines. Similar findings were reported by Rappolt (1997). Such results give ideas for suggestions on how to develop guidelines further. Effort should be aimed at developing guidelines that are easier to use and easier to find when needed; guideline programmes may offer opportunities to improve availability (Jousimaa *et al*. 1998). Resources and abilities needed by health care professionals and organisations should be developed to match new guidelines.

The 14-item common survey instrument developed here, the Attitudes towards Guidelines Scale (AGS) can be a useful tool for comparing organisations and occupational groups in primary care to find out the level of various perceived barriers and facilitators in different guideline implementation projects. The validity of the AGS needs to be determined separately in other countries and in hospital settings.

KEY MESSAGES

- Important factors behind general positive or negative attitudes towards guidelines are the usefulness, reliability, practicality and availability of the guidelines. Also the overall individual, team and organisational competence to follow the recommended procedures seem to be vital.
- All these factors could be measured with the Attitudes Towards Guidelines Scale (AGS), a 14-item instrument developed within CPP.
- Positive attitudes towards guidelines were associated with the reported use of clinical guidelines.
- The validity of the instrument for countries and settings other than Finnish primary care needs to be determined separately.

References

Anderson JC, Gerbing DW. Structural equation modelling in practice: A review and recommended two-step approach. *Psychological Bulletin* 1988; 103: 411-23.

Cohen J, Cohen P. *Applied multiple regression/correlation analysis for the behavioral sciences*. 2nd ed. Hillsdale: Erlbaum, 1983.

Eccles M, Clapp Z, Grimshaw J, Adams CP, Higgins B, Purves I, Russell I. Developing valid guidelines: methodological and procedural issues from North of England Evidence Based Guideline Development Project. *Quality in Health Care* 1996; 5(1): 44-50.

Eddy DM. Designing a practice policy. Standards, guidelines and options. *JAMA* 1990; 263(22): 3077-84.

Feder G, Eccles M, Grol R, Griffiths C, Grimshaw J. Using clinical guidelines. *British Medical Journal* 1999; 318: 728-730.

Field MJ, Lohr KN. (editors). Institute of Medicine. *Clinical practice guidelines: Directions for a new program.* Washington: National Academy Press, 1990.

Grimshaw J, Eccles M. Clinical practice guidelines. Pp. 110-122 in: Silagy C, Haines A. (eds.) *Evidence based practice in primary care.* London: BMJ Books, 1998.

Grimshaw J, Russell I. Effect of clinical guidelines on medical practice: a systematic review of rigorous valuations. *The Lancet* 1993; 342: 1317-22.

Grol R. Thomas S, Roberts R. Development and implication of guidelines for family practice: Lessons from the Netherlands. *J Fam Pract* 1995; 40(5): 435-39.

Grol R. National standard setting for quality of care in general practice: attitudes of general practitioners and response to a set of standards. *Br J Gen Pract* 1990; 40: 361-64.

Hayduk LA. *Structural Equation Modeling with Lisrel: Essentials and Advances.* Baltimore: The John Hopkins Univ. Press, 1989.

Hellbruck RP. Medical guidelines: A valid and reliable management tool. *International Journal of Health Planning and Management* 1997; 12: 51-62.

Jousimaa J, Kunnamo I, Mäkelä M. Physicians' patterns of using a computerized collection of guidelines for primary care. *International Journal of Technology Assessment in Health Care* 1998; 14(3): 484-493.

Jöreskog KG, Sörbom D. *Lisrel 7 User's Reference Guide.* Scientific Software, 1993.

Kanouse DE, Jacoby I. When does information change practitioners' behavior? *International Journal of Technology Assessment in Health Care* 1988; 4: 27-33.

Mansfield CD. Attitudes and behaviors towards clinical guidelines: the clinicians' perspective. *Quality in Health Care* 1995; 4(4): 250-55.

Rappolt SG. Clinical guidelines and the fate of medical autonomy in Ontario. *Soc Sci Med* 1997; 44(7): 977-987.

Varonen H, Mäkelä M. Practice guidelines in Finland: availability and quality. *Quality in Health Care* 1997; 6: 75-79.

Annex 1: Statistical Appendix

The AGS instrument was tested in four stages. First, we analysed the structure of the whole scale using an exploratory factor analysis (maximum likelihood) and varimax rotation (Sample 1). After the first analysis, we dropped items which did not fit any of the tested structures.

Second, to explicate the final structure, to get structural estimates of how well the items measure the factor structure and to reduce the number of items, we performed a confirmatory factor analysis with LISREL 8 (Jöreskog *et al.* 1993) for Samples 1 and 2. In the confirmatory factor analysis, a polychoric correlation matrix, computed by the PRELIS 2 programme, was used as input for the LISREL program (Anderson *et al.* 1988). Each subset of measured items was allowed to load only on its corresponding latent variable. No correlation errors either within or across sets of items were allowed in the model, because they could not be predicted a priori based on the theory.

A number of indices were used to evaluate the accuracy of the factor structure. When an overall goodness of fit test is used, the null hypothesis is that the data provide a good fit to the theoretical model. In principle, a non-significant value in the X^2 test would signify that the data provide a good fit to the model. A problem with the X^2 test is, however, that with large samples any theoretical model is usually rejected (Hayduk 1989). To avoid this, several other statistics can be used to support the evaluation of the adequacy of a model. The degrees of freedom for X^2 are the function of the number of observed variables analysed and the total number of independent parameters estimated. The X^2 measure is sensitive to sample size and very sensitive to departures from multivariate normality of the observed variables (Jöreskog *et al.* 1993).

Root mean square residuals (RMSR) are indicative of the discrepancies between the observed and predictive relations. The smaller the RMSR, the better the fit. The adjusted goodness of fit test (AGFI) also indicates the fit: the greater the value, the better the model fit. The Bentler and Bonett normed fit index (BBI) and the Tucker-Lewis index (TLI) are indices of how well the theory specified model fits the data relative to other competing models. The value of these indices (AGFI, BBI and TLI) ranges between 1 and 0, with large values indicating better fit. Of the indices mentioned above, the TLI is the least dependent on sample size. All the indices were used in this study.

Third, we calculated the internal consistency of the questionnaire and the final scales (Cronbach's alpha) produced in stages 1 and 2. And finally, regression analysis (Cohen & Cohen, 1983) was used to test the predictive validity of the final scale.

Testing of the 14-item scale

The full model of six subscales, including 27 items from the original scale, did not fit the data due to the correlating errors of the items [$X^2(309)=1057.56$ ($p=0.000$)]. Items were excluded one by one following modification indices provided by LISREL. The first acceptable measurement model, in which the number of items was reduced to 14 and one factor was split into two, is presented in Table 8.1 (page 157). The X^2 tests for the model [$X^2(56)=103.97$ ($p=0.0001$)] did not indicate a very good fit of the model to the data. The other indices, on the other hand, supported the conclusion that the model provides a reasonably good fit to the data (RMSR=0.056, AGFI=0.85, BBI=0.87, TLI=0.89). The LISREL programme itself suggests changes in the model that may improve the fit. However, none of the suggested changes significantly improved the fit. The factor loadings of the model are presented in Table 8.1. The factor loadings ranged from 0.34 to 0.97. These loadings essentially indicate that most of the observed items measure the latent constructs at least acceptably. The most serious exception is the last item (on availability) with the lowest loading.

The final structure was tested also for sample 2. The X^2 tests for the model with sample 2 [$X^2(56)=182.56$ ($p=0.0001$)] did not indicate that the model provides a very good fit to the data. Again the other indices, on the other hand, supported the conclusion that the model provides a reasonably good fit to these other data (RMSR=0.049, AGFI=0.90, BBI=0.89, TLI=0.90). The loadings ranged from 0.32 to 0.83 in sample 2.

Annex 2: Items and reliability (Chronbach's alpha) of the subscales of the 27-item Attitudes Towards Guidelines Scale

1. Usefulness of guidelines – 6 items, 0.86
   - Guidelines are useful as educational tools.
   - Guidelines are a convenient source of advice.
   - Guidelines are useful in clinical work.
   - Guidelines can facilitate communication with patients and families.
   - Guidelines can improve the quality of health care.
   - Guidelines can facilitate team-oriented health-care delivery.

2. Reliability of guidelines – 4 items, 0.88
   - Guidelines are based on dependable information.
   - Guidelines are based on scientific evidence.
   - Guidelines are made by experts.
   - Guidelines can be trusted.
3. Lack of individual competence – 4 items, 0.71
   - I do not have enough information about guidelines to implement them.
   - My occupational competence is insufficient for adopting the latest guidelines.
   - Most of our team members have disapproving attitudes about guidelines.
   - Our team is not unanimous enough to accept one any single guideline.

4. Lack of organisational competence – 4 items, 0.72
   - We do not have enough staff to implement guidelines.
   - Guidelines are not valued in our organisation.
   - To implement guidelines is too expensive for us.
   - We do not have the right kind of staff for implementing guidelines.

5. Impracticality – 4 items, 0.65
   - Guidelines challenge the autonomy of care providers.
   - Guidelines oversimplify medical practice.
   - Guidelines are impractical.
   - Long guidelines with too many pages are inconvenient in real situations.

6. Availability – 5 items, 0.75
   The availability of guidelines at our premises is good.
   Guidelines are difficult to find if needed. (Reversed)
   Our central hospital provides local treatment programmes based on guidelines.
   I have not seen any guidelines in our health care unit. (Reversed)
   Guidelines are used in my unit for quality review.

Annex 3. Means, standard deviations and correlations (Pearson's) between AGS subscales

Sample 1

	M	*SD*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
(1) General attitude	5.4	1.1	-						
(2) Usefulness	5.3	0.9	.51	-					
(3) Reliability	5.4	1.1	.38	.52	-				
(4) Lack of individual or team comp.	3.1	1.3	-.24	-.30	-.34	-			
(5) Lack of organisational comp.	2.9	1.1	-.22	-.25	-.24	.50	-		
(6) Impracticality	3.5	1.0	-.27	-.35	-.34	.27	.20	-	
(7) Availability	4.3	1.4	.23	.25	.23	-.20	-.31	-.31	-

Sample 2

	M	*SD*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
(1) General attitude	5.5	1.0	-						
(2) Usefulness	5.7	0.9	.62	-					
(3) Reliability	5.6	1.0	.43	.48	-				
(4) Lack of individual or team comp.	2.5	1.1	-.14	-.14	-.21	-			
(5) Lack of organisational comp.	5.2	1.1	-.15	-.10	-.11	.20	-		
(6) Impracticality	3.1	1.2	-.31	-.30	-.28	.27	.34	-	
(7) Availability	3.7	1.6	.26	.20	.22	-.25	.37	-.33	-

Annex 4. Attitudes towards guidelines among different occupational groups.

Means (M) and standard deviations (SD)

	Physicians and dentists M(SD)	*Registered nurses M(SD)*	*Practical nurses M(SD)*
Sample 1			
General attitude towards guidelines	5.4(1.2)	5.3(1.2)	5.4(1.0)
Usefulness of guidelines	5.3(0.7)	5.4(0.9)	5.4(0.9)
Reliability of guidelines	5.6(1.2)	5.4(1.1)	5.4(1.0)
Lack of individual competence to use guidelines	2.8(1.1)	3.0(1.4)	3.3(1.2)
Lack of organisational competence to use guidelines	2.7(1.3)	2.9(1.1)	2.9(1.1)
Impracticality of guidelines	3.4(0.9)	3.3(1.1)	3.5(1.1)
Availability of guidelines	4.3(1.4)	4.2(1.5)	4.3(1.4)
Sample 2			
General attitude towards guidelines	5.7(1.1)	5.7(0.9)	5.3(1.0)
Usefulness of guidelines	5.8(0.8)	5.5(0.9)	5.6(0.7)
Reliability of guidelines	5.8(0.8)	5.8(0.8)	5.4(0.8)
Lack of individual competence to use guidelines	2.1(0.7)	2.3(1.0)	2.6(1.0)
Lack of organisational competence to use guidelines	4.7(0.9)	4.9(0.8)	5.2(1.0)
Impracticality of guidelines	3.0(1.2)	2.9(1.2)	3.3(1.1)
Availability of guidelines	3.2(1.3)	3.5(1.5)	3.7(1.5)

Annex 5: Regression results for effect of AGS subscales on the use of guidelines, for samples 1 and 2

Independent variables

	estimate	*F-value*
Sample 1		
Lack of individual or team competence to use guidelines	-0.17	11.80**
Availability of guidelines	0.47	35.08***
R^2 0.29		
F 14.01***		
Sample 2		
Lack of individual or team competence to use guidelines	-0.13	4.49*
Reliability of guidelines	0.15	3.97*
Availability of guidelines	0.71	217.0***
R^2 0.45		
F 37.09***		

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$

Chapter 9

Researching perceived barriers and facilitators to implementation: a coded review of studies

H. Claudia Pagliari, James P. Kahan

Aims of the chapter

- To provide a bibliographic review of research on perceived or attitudinal barriers and facilitators to implementation.
- To identify studies which have used structured instruments to investigate these variables.
- To categorise the studies listed in the bibliography in such a way as to facilitate the identification of appropriate instrument for future projects.

Introduction and aims

As has been discussed in several of the preceding chapters, clinical guidelines are often not followed by the health professionals whose practice they are designed to inform. Several authors on the subject have suggested that guideline adoption and use will be greatly facilitated by taking account of barriers to change when designing implementation. It is for this reason that appropriate methods for identifying and measuring such barriers must be found. Wensing *et al.* describe a number of suitable methods in chapter 6 of this monograph, along with relevant theories arising from the literature. Re-visiting these methods and theories is not the purpose of this chapter. Our primary aim here is to provide a categorised bibliography of studies that have investigated perceived facilitators and barriers using (primarily) structured methods. This, we hope, will be a useful resource for researchers planning new implementation

projects. The bibliography presented here is a direct outcome of our work to develop the CPP 'Perceptions Library' – an Internet website containing examples of questionnaire items pertinent to this issue. The studies reported in this chapter and the items reproduced on the CPP website are particularly concerned with assessing individuals' judgements or opinions about aspects of a health technology or change, or of their environment, which they believe to be a help or a hindrance to implementation. Barriers rooted in people's subjective beliefs or attitudes (e.g. belief that one's organisation does not support the change) can be just as influential in determining their behaviour as genuine barriers (e.g. resource constraints). Moreover, they can be difficult to anticipate since they may not relate to directly observable or quantifiable features of the environment. It is for this reason that the 'Perceptions Library' was envisaged.

Literature review methods used to identify studies containing perceptual items

We conducted a comprehensive literature search using several computerised databases (see Box 9.1) and soliciting information on other pertinent studies from colleagues in the field. We also contacted all CPP partners in order to obtain details of unpublished studies or those published in languages other than English.

Box 9.1
Summary of Search Procedures

- **Databases searched:** Medline, Embase, Cinahl, Psychlit. 1975 – 1998 (Prior to 1975 the search term 'practice guide-line' did not produce any 'hits' on Medline, reflecting the relative youth of this innovation. For this reason it was not thought worthwhile to extend the search further back in time.)
- **Search terms:** The following search terms were used singly, or in various combinations, to reveal the maximum number of pertinent abstracts: *Perceived/perception, barrier/barriers, facilitator/facilitators, practice guideline/ guideline, implement/implementation, professional behaviour change, attitude/attitudes, survey/ questionnaire.*
- **Sifting process:** Titles and abstracts were scanned. Abstracts appearing to meet the aims of the study were selected and the full-text article examined for relevant items. Where suitable items were alluded to but not reproduced in the article, authors were contacted for further details.

We focused our literature search on studies dealing with the implementation of new health technologies or examining professional behaviour change amongst health care workers. Although we examined a number of review articles, we paid particular attention to studies which had used structured survey/interview methods to investigate perceived facilitators and barriers (including attitudinal barriers and explicit barriers). Although our search focused on questionnaire/survey research, studies using less structured methods were included if it was felt that their methods or findings could be used to produce structured items for future research (e.g. semi-structured interview items, coding frames derived from qualitative analysis). Box 9.2 summarises the broad scope of the search.

BOX 9.2
THE SCOPE OF THE SEARCH

- Behaviour of ***health professionals*** (primary and secondary care doctors and nurses, other health care professionals) Not patients or managers.
- ***Implementation*** of new technologies (e.g. guidelines, drug prescribing), general strategies (e.g. evidence based medicine, using research), other ***professional behaviour change*** (e.g. changes in management).
- ***'Attitudes', perceived obstacles, perceived facilitators*** (Not actual practice or knowledge, although these were regularly investigated in the studies reviewed and are assigned category labels in table 9.3).
- ***Structured survey/interview methods*** (Some of the studies also used unstructured methods).

The products of our search are shown in the bibliography appearing at the end of this article.

CATEGORISATION OF STUDIES

Although many items extracted from the articles listed in the bibliography may be found on the CPP website, a wider range will be available by directly accessing the primary articles. To assist researchers in identifying suitable literature sources and methods, we have categorised the articles mentioned in our bibliography using the scheme shown in Box 9.3 (page 172).

Box 9.3
Categorisation scheme

1. Focus
1a. Clinical practice guidelines
1b. Evidence-based medicine/utilisation of research
1c. Adoption of new organisational methods/procedures (e.g. managed care, quality assurance)
1d. Specific practices (e.g. breast cancer screening, test ordering, 'green' prescribing)
1e. General practices (e.g. caring for mentally disabled adults, managing obesity)
1f. Professional role behaviours (e.g. emergency nurse practitioner role)
1g. Educational measures
1h. Other (e.g. needs of remote health care professionals)

2. Targets
2a. Primary care doctors (GPs, family physicians)
2b. Secondary care physicians and specialists (e.g. cardiologists, paediatricians, obstetricians/ gynaecologists)
2c. Other specialities: mental health, dentistry
2d. Nurses (general)
2e. Nurse practitioners & midwives
2f. Professions allied to medicine (e.g. dieticians, physiotherapists)
2g. Educators or managers (involved in training or supervising others)
2h. Teams (explicit multiprofessional team target)
2z. Unspecified ("physicians") or multiple groups ("health care providers")

3. Methods
3a. Postal or other mass-distributed questionnaire/survey
Type: *3.a.i closed, 3.a.ii open, 3.a.iii mixed*
3b. Interview (face-to-face or telephone)
Type: *3.b.i structured, 3.b.ii semi-structured (or mixed), 3.b.iii open (in depth)*
3c. Focus group
3d. Reference to cases *(critical incident, constructed scenario, diaries)*
3e. Document analysis *(e.g. practice patterns, objective measures of compliance with guideline recommendations, other audit measures)*
3f. Observation
3z. Other method
3x. unknown

4. Object
4a. "KAP" *(knowledge [cognitions], attitudes [beliefs, opinions], practice [behaviour, intention])*
4b. 'Attitudes' *(about guidelines or other innovation. E.g. applicability, helpfulness)*
4c. Familiarity *(with guideline/innovation)*
4d. Barriers or Facilitators
4d.i General
4d.ii Legal
4d.iii Social/organisational
4d.iv. Professional/role (e.g. threat to autonomy)
4d.v Time
4d.vi Resource (economic, access)
4e. Propensity to change, self-efficacy
4f. Perceived need, motivation
4g. Provider characteristics ("demographics")
4h. Strategies for change
4i. Actual change.

R. Review articles
(not containing original research, but dealing with key concepts and measures)

Table 9.1 gives examples of studies associated with several categories.

Table 9.1. Examples of study summaries, with attached classifications according to the scheme shown in Box 9.3.

Christakis DA, Rivara FP. (1998) Pediatricians' awareness of and attitudes about four clinical practice guidelines. Pediatrics 101, 5, 825-830.	Surveyed US pediatricians about their knowledge and impressions of four well-publicised clinical practice guidelines. Research questions: What percentage of practicing pediatricians is aware of these guidelines? How helpful do they find them? What are practitioners' perceived limitations of these guidelines? Have these guidelines affected provider behaviour? Are there features of a provider's training or practice that are associated with changing practice as a result of guidelines?	1a. 2b. 3a. 4b. 4c. 4g. (4a)
Ely JW, Goerdt CJ, Bergus GR, West CP, Dawson JD, Doebbeling BN. (1998) The effect of physician characteristics on compliance with adult preventive care guidelines. Family Medicine 30, 1, 34-39.	Aim: To identify physician characteristics and attitudes related to self-reported compliance with adult prevention guidelines. Methods: Mailed questionnaire to family practice and internal medicine residents and faculty at the University of Iowa. Details of instrument: 78 items in seven categories: physician characteristics, history-taking practices, counselling practices, self perceived effectiveness in changing patient behaviour, beliefs about preventive care, knowledge about preventive care, and perceived barriers to the delivery of preventive care. (e.g. lack of time, lack of reminder systems, attitudes about preventive care etc.).	1a. 2a. 3a. (4a) 4b, 4d, 4e. 4g. 4i.
Bernat JL, Ringel SP, Vickrey BG, Keran C. (1997) Attitudes of US neurologists concerning the ethical dimensions of managed care. Neurology 49 1, 4-13.	Surveyed attitudes of US neurologists about the ethical dimensions of managed care by administering a written instrument containing paradigmatic cases portraying conflicts of physicians, patients and managed care organizations. After each case, assessed neurologists' attitudes by asking them their degree of agreement with a series of statements. Found that neurologists 1) generally were willing to follow clinical practice guidelines if they were created by medical societies, 2) experienced frequent conflicts of interest or conflicts of obligation in the care of their MCO patients, 3) feared legal ramifications of their clinical decisions on MCO patients, 3) were unwilling to employ deception or gaming to achieve what they perceived to be good patient care, 5) believed that their professional prerogatives and autonomy were under attack by MCOs and 6) felt that the good of their patients should not be sacrificed for the good of society.	1c. 2b. 3a. 3d. 4b. 4dii. 4diii.
Carroll DL, Greenwood R, Lynch KE, Sullivan JK, Ready CH, Fitzmaurice JB. (1997) Barriers and facilitators to the utilization of nursing research. Clinical Nurse Specialist 11, 5, 207-212	Aim: "to explore the nurse's perception of the barriers and facilitators to using research findings in nursing practice." "A survey methodology was used and a sample of 356 practicing (US) nurses responded. Data were collected using a scale that rated the barriers and facilitators to using research findings in nursing practice." "The greatest barriers were insufficient time on the job to implement new ideas, lack of knowledge of nursing research findings, and inaccessibility of relevant literature."	1b. 2d. 3a. 4d.

Costanza ME, Stoddard AM, Zapka JG, Gaw VP, Barth R. (1992) Physician compliance with mammography screening guidelines: barriers and enhancers. Journal of the American Board of Family Practice 5, 2, 143-152	Examined the attitudes, beliefs and practices with regard to breast cancer screening, as self-reported by New England primary care physicians. Mailed survey. Included questions on attitudes and beliefs about breast cancer screening, as well as questions about perceived barriers and actual screening practices. "Strongly associated with ordering annual mammograms were beliefs in the benefit of mammography and the perception of community consensus regarding breast cancer screening. A strong positive association of practising in a group setting and mammography compliance was noted.... The three most important determinants of annual screening suggest ways to improve physician compliance: Improve physician attitudes about the benefits of mammography, build further on the medical community's consensus regarding the appropriateness and importance of annual guidelines, target the poorest compliers with special messages or programs."	1d. 2a. 3a. 4a.

Summaries of the articles included in our bibliography (such as shown in Table 9.1) can be found in a separate publication (Pagliari & Kahan, 1999).[1)]

Summary

Our literature search has revealed a large number of studies that have investigated perceived or attitudinal barriers to the implementation of health care technologies. Relatively few of these have focused specifically on clinical guidelines. For example, of the 135 titles listed in our bibliography, only 37 mention guidelines. Many have examined comparable innovations and changes, however, such as those dealing with barriers to research utilisation and evidence-based medicine. Across those studies which have employed structured measures, there has been considerable variation in the instruments used. No single instrument can be recommended as a general measure of perceived or attitudinal barriers or facilitators, and the investigator's choice of measurement tool should be guided by the specific focus and context of the research questions. We hope that future implementation researchers will find our list of classified references helpful for the purposes of targeting their literature review and identifying suitable instruments to be used in or adapted for their studies. We suggest that this database be used in conjunc-tion with the CPP Perceptions Library and the list of annotated abstracts mentioned earlier (Pagliari & Kahan 1999).

1) Pagliari, HC & Kahan, JP (1999). "Changing Professional Practice in Health Care: Annotated Bibliography of studies of Perception in Guidelines Implementation". Leiden: RAND Europe Report 99.008.

Key Messages

- A considerable number of studies have investigated perceived or attitudinal barriers to implementation, using structured methods.
- No single instrument may be recommended as a global measure. Selection of an appropriate instrument will depend upon the research questions and context, including the technology itself (focus), the personnel involved (targets) and the dimensions of change of interest (objects).
- This categorised bibliography may be helpful to researchers as a means of focusing their review of the literature and the available instruments to those which are most appropriate to their specific objectives.

Categorised bibliography

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
1. **Allen RE, Read J. (1997)** Integrated mental health care: practitioners' perspectives. Australian and New Zealand Journal of Psychiatry 31, 4, 496-503.	c	c	ai, biii	b, di, dii, dv, dvi, i
2. **Allery LA, Owen P, Robling MR. (1997)** Why general practitioners and consultants change their clinical practice: a critical incident study. British Medical Journal 314, 870-874.	g	a, b	bii, d	h, i
3. **Alston, RP. (1997)** Guidelines and cardiac anaesthetists. Not in my backyard. Anaesthesia 52, 328-331.	a	b	aiii	b, di, c
4. **Anderson AL, Gilliss CL, Yoder L. (1996)** Practice environment for nurse practitioners in California. Identifying barriers. Western Journal of Medicine 165, 4, 209-214.	f	e	aiii	dii, diii
5. **Armstrong SD, Reyborn H, Jones R. (1996)** A study of general practitioners' reasons for changing their prescribing behaviour. BMJ 312, 949-952.	d	a	biii	i
6. **Armstrong D, Fry J, Armstrong P. (1994)** General practitioners' views of clinical guidelines for the management of asthma. International Journal of Quality in Health Care, 6, 199-202.	a	a, b	Ai	b
7. **Bakemeier RF, Krebs LU, Murphy JR, Shen Z, Ryals T. (1995)** Attitudes of Colorado health professionals toward breast and cervical cancer screening in Hispanic women. Journal of the National Cancer Institute. Monographs 18, 95-100.	d	z	ai, c	c, di
8. **Banazak DA. (1996)** Late-life depression in primary care. How well are we doing? Journal of General Internal Medicine 11, 1, 163-167.	d	a	ai	a
9. **Barta KM (1995)** Information-seeking, research utilization, and barriers to research utilization of pediatric nurse-educators. Journal of Professional Nursing 11, 1, 49-57.	b	g	ai	di
10. **Bernat JL, Ringel SP, Vickrey BG, Keran C. (1997)** Attitudes of US neurologists concerning the ethical dimensions of managed care. Neurology 49, 1, 4-13.	h	b	ai, d	b

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
11. **Berrow D, Humphrey C, Hayward J. (1997)** Understanding the relation between research and clinical policy: a study of clinicians' views. Quality in Health Care 6, 181-186.	b	h	aii, bii, e, f	b, di, f, g
12. **Bookbinder M, Coyle N, Kiss M, Goldstein ML, Holritz K, Thaler H, Gianella A, Derby S, Brown M, Racolin A, Ho MN, Portenoy RK. (1996)** Implementing national standards for cancer pain management: program model and evaluation. Journal of Pain and Symptom Management 12, 6, 334-347.	d	d	c	a
13. **Bull FC, Schipper EC, Jamrozik K, Blanksby BA. (1997)** How can and do doctors promote physical activity? Preventive Medicine 26, 6, 866-873.	d	a	[x]	a
14. **Carroll DL, Greenwood R, Lynch KE, Sullivan JK, Ready CH, Fitzmaurice JB. (1997)** Barriers and facilitators to the utilization of nursing research. Clinical Nurse Specialist 11, 5, 207-212.	b	d	ai	di
15. **Carter AO, Battista RN, Hodge MJ, Lewis S, Basinski A, Davis D. (1995)** Report on activities and attitudes of organizations active in the clinical practice guidelines field. Canadian Medical Association Journal 153, 7, 901-907.	a	z	aiii	b, di, h
16. **Christakis DA, Rivara FP. (1998)** Pediatricians' awareness of an attitudes about four clinical practice guidelines. Pediatrics 101, 5, 825-830.	a	b	ai	a
17. **Costanza ME, Stoddard AM, Zapka JG, Gaw VP, Barth R. (1992)** Physician compliance with mammography screening guidelines: barriers and enhancers. Journal of the American Board of Family Practice, 5, 2, 143-152.	d	a	ai	a
18. **Craven MA, Cohen M, Campbell D,** Williams J, Kates N. (1997) Mental health practices of Ontario family physicians: a study using qualitative methodology. Canadian Journal of Psychiatry – Revue Canadienne de Psychiatrie 42, 9, 943-949.	d	a	c	a

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
19. **Davis D, Thomson MA, Oxman A, Haynes R, Bryan MD. (1995)** Changing physician performance: A systematic review of the effect of continuing medical education strategies. JAMA 274, 9, 700-705.	REVIEW			
20. **Department of Health and Human Services. (1994)** The NIH Consensus Development Program: Dissemination of Findings through Medical School Continuing Education. DHHS publication no. OEI-019101706 Feb. 1994.	g	g	ai, biii	a
21. **Dunn V, Crichton N, Roe B, Seers K, Williams K. (1997)** Using research for practice: a UK experience of the BARRIERS scale. Journal of Advanced Nursing 26, 6, 1203-1210.	b	d	ai	di
22. **Elovainio M, Mäkelä M, Sinervo T, Kahan J. (1998)** The effects of job characteristics, team climate and attitudes towards guidelines on the use of clinical guidelines (Quality in Health Care – Submitted).	a	z	ai	b, di, f
23. **Ely JW, Goerdt CJ, Bergus GR, West CP, Dawson JD, Doebbeling BN. (1998)** The effect of physician characteristics on compliance with adult preventive care guidelines. Family Medicine 30, 1, 34-39.	a	a	ai	a, di, g
24. **EUR-ASSESS PROJECT SUBGROUP REPORT ON DISSEMINATION AND IMPACT. (1997)** International Journal of Technology Assessment in Health Care, 13, 2, 220-286.	REVIEW			
25. **Feldman EL, Jaffe A, Galambos N, Robbins A, Kelly RB, Froom J. (1998)** Clinical practice guidelines on depression: awareness, attitudes, and content knowledge among family physicians in New York. Archives of Family Medicine 7, 1; 58-62.	a	a	ai	a, g
26. **Felton T, Lister G. (1996)** Consider the evidence. The NHS on the move towards evidence based medicine. Uxbridge: Coopers and Lybrand.	b	a, b	a.iii	a, d, f
27. **Ferrier B, Woodward CA, Cohen M, Williams AP. (1996)** Clinical practice guidelines. New-to-practice family physicians' attitudes. Canadian Family Physician 42, 463-468.	a	A	ai	b
28. **Firth-Cozens J. (1997)** Healthy promotion: changing behaviour towards evidence-based health care. Quality in Health Care, 6, 205-211.	REVIEW			

Source	*11. Focus*	*2. Targets*	*3. Methods*	*4. Object*
29. **Fleuren MAH. (1997)** Managing (imminent) miscarriage in primary health care. An evaluation of the (imminent) miscarriage guideline of the Dutch College of General Practitioners. PhD Thesis. Free University of Amsterdam. (5 studies in total. Studies 1, 2, and 3 most pertinent to 'perceptions' library).	Study 1: a Study 2: a Study 3: a	Study 1: a, e Study 2: a Study 3: e	Study 1: ai Study 2: bii, e Study 3: bii, e	Study 1: a, d Study 2: b, d Study 3: b, d
30. **Flottorp S, Oxman A, Bjorndal A. (1998)** Shooting in the dark? Barriers to appropriate use of laboratory tests. Manuscript: Oslo: Health Services Research Unit, National Institute for Public Health.	d	a	ai, c	di
31. **Friedman LC, Plon SE, Cooper HP, Weinberg AD. (1997)** Cancer genetics – survey of primary care physicians; attitudes and practices. Journal of Cancer Education 12, 4, 199-203.	d	a	ai	b, di
32. **Funk SG, Champagne MT, Wiese RS, Tornquist EM (1991a)** BARRIERS: the barriers to research utilization scale. Applied Nursing Research 4, 1, 39-45.	b	d	a.iii	d.i, g
33. **Funk SG, Champagne MT, Wiese RS, Tornquist EM (1991b)** Barriers to using research findings in practice: the clinician's perspective. Applied Nursing Research 4, 2, 90-95.	b	d	a.iii	d.i, g
34. **Funk SG, Tornquist EM, Champagne MT. (1995)** Barriers and facilitators of research utilization: an integrative review. Nursing Clinics of North America 30, 3; 389-407.	REVIEW			
35. **Gilbert L, Brand AA, Rudolph MJ. (1995)** Private dental practitioners – perceptions of needs and barriers to dental care. Journal of the Dental Association of South Africa 50, 1, 5-7.	e	c	[x]	di
36. **Glasser M, Gravdal JA. (1997)** Assessment and treatment of geriatric depression in primary care settings. Archives of Family Medicine 6, 5, 433-438.	d	a	ai	b, di (f), g
37. **Glanz K (1979)** Strategies for nutritional counseling. Dieticians' attitudes and practice. Journal of the American Dietetic Association 74, 4, 431-437.	d	f	ai	a

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
38. **Grant AM, Niyonsenga T, Dion I, Delisle E, Xhignesse M, Bernier R. (1998)** Cardiovascular disease: Physician attitudes toward prevention and treatment. Canadian Family Physician 44, 780-787.	a	a	bi	a
39. **Greco PJ, Eisenberg JM. (1993)** Changing Physicians' Practices. New England Journal of Medicine 329, 17, 1271-1273.	REVIEW			
40. **Griffin MA. (1992**) An exploratory study of the influencing factors of prescriptive authority for nurse practitioners. Ed.D Thesis. Seton Hall University, School of Education.	d	d	ai	di, g
41. **Grilli R. (1996)** Developing recommendations to promote the uptake of research information in clinical practice. Therapie 51, 3, 265-268.	REVIEW			
42. **Grilli R, Penna A, Zola P, Liberati A. (1996)** Physicians' view of practice guidelines. A survey of Italian physicians. Social Science & Medicine, 43, 8, 1283-1287.	d	b	ai	b
43. **Grilli R, Lomas J. (1994)** Evaluating the message: The relationship between compliance rate and the subject of a practice guideline. Medical Care, 32, 3, 202-213	REVIEW			
44. **Grilli R, Apolone G, Marsoni S, Nicolucci A, Zola P, Liberati A. (1991)** The impact of patient management guidelines on the care of breast, colorectal and ovarian cancer patients in Italy. Medical Care, 29, 1; 50-63	a	b	ai	a
45. **Grol R. (1990)** National standard setting for quality of care in general practice: attitudes of general practitioners and response to a set of standards. British Journal of General Practice 40, 361-364.	c	a	ai	b, c, di
46. **Grol R. (1992)** Implementing guidelines in general practice care. Quality in Health Care, 1: 184-191.	REVIEW			
47. **Grol R. et al (1996)** Effectiveness of methods and interventions to implement guidelines and change professional behaviour. Nijmegen; Centre for Quality of Care Research. (In Dutch).	d	a	[x]	b, g
48. **Grol R. (1997)** Beliefs and evidence in changing clinical practice. British Medical Journal 315, 418-421.	REVIEW			

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
49. **Grol R, Dalhujisen J, Thomas S, in 't Veld C,** Rutten G, Mokkink H. (1998) Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. BMJ 317, 858-861.	REVIEW			
50. **Grol R, Wensing M. (1995)** Implementation of quality assurance and medical audit: general practitioners' perceived obstacles and requirements. British Journal of General Practice 45, 547-552.	c, d	a	bii	a, d, g, f
51. **Gupta L, Ward JE, Hayward RSA. (1997)** Clinical practice guidelines in general practice: a national survey of recall, attitudes and impact. Medical Journal of Australia 166, 2, 69-92.	a	a	aiii	a, d
52. **Hamilton FA, Hill FJ, Holloway PJ. (1997)** An investigation of dento-alvular trauma and its treatment in an adolescent population. Part 2: Dentists' knowledge of management methods and their perception of barriers to providing care. British Dental Journal 182, 4, 129-133.	d	c	ai, bi	c, di
53. **Hancock R, Bender P, Dayhoff N, Nyhuis A. (1996)** Factors associated with nursing interventions to reduce incontinence in hospitalized older adults. Urologic Nursing 16, 3, 79-85.	d	d	ai, d	di
54. **Hassanein RS. (1988)** Barriers to research in allied health. Journal of Allied Health 17, 3, 175-187.	b	z	aiii	di
55. **Hayden ML, Davies LR, Clore ER. (1982)** Facilitators and inhibitors of the emergency nurse practitioner role. Nursing Research 31, 5, 294-299.	e	e	aiii	di
56. **Hayward RSA, Guyatt GH, Moore KA, McKibbon A, Carter AO. (1997)** Canadian physicians' attitudes about and preferences regarding clinical practice guidelines. Canadian Medical Association Journal 156, 12, 1715-1723.	a, b	a, b, c	ai	a, di, g
57. **Helwig A, Bower D, Wolff M, Guse C. (1998)** Residents find clinical practice guidelines valuable as educational and clinical tools. Family Medicine 30, 6, 431-435.	a	a	aiii	a
58. **Henry RC, Ogle KS, Snellman LA (1987)** Preventive medicine: physician practices, beliefs, and perceived barriers for implementation. Family Medicine 19, 2, 110-113.	e	a	ai	a

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
59. **Herman CJ, Lengerich EJ, Stoodt G. (1996)** Variation in recommendations for breast and cervical cancer screening among primary care physicians in North Carolina, 1991. Southern Medical Journal 89, 6, 583-589.	d	a	ai	a
60. **Hicks C, Hennessy D, Cooper J, Barwell F. (1996)** Investigating attitudes to research in primary health care teams. Journal of Advanced Nursing 24, 5, 1033-1041.	b	a	bii	di
61. **Hiddink GJ, Hautvast JG, van Woerkum CM, Fieren CJ, van 't Hof MA. (1995)** Nutrition guidance by primary care physicians: perceived barriers and low involvement. European Journal of Clinical Nutrition 49, 11, 842-851.	d	a	ai	di
62. **Hughes RM. (1993)** Exploratory study of factors influencing nurses' participation in continuing education in public and private health care settings. Georgia State University, PhD thesis.	g	d	ai	di
63. **Hulscher MEJL, Van Drenth BB, Mokkink, HGA, Van der Wouden, JC, Grol RPTM. (1997)** Barriers to preventive care in general practice: the role of organizational and attitudinal factors. British Journal of General Practice, 47, 711-714.	e	a	ai, f	b, di, e, g
64. **Hutchison BG, Abelson J, Woodward CA, Norman G. (1996)** Preventive care and barriers to effective prevention. How do family physicians see it? Canadian Family Physician 42, 1693-1700.	e	a	ai	di
65. **Inouye J, Kristopatis R, Stone E, Pelter M, Sandhu M, Weingarten, S. (1998)** Physicians' changing attitudes toward guidelines. Journal of General Internal Medicine 13, 5, 324-326.	e	a	ai	b, di
66. **James PA, Cowan TM, Graham RP, Majeroni BA. (1997)** Family physicians' attitudes about and use of clinical practice guidelines. Journal of family practice 45, 4, 341-347.	c	a	ai	b
67. **Kanouse DE, Jacoby I . (1988)** When does information change practitioners' behaviour? International Journal of Technology Assessment in Health Care 4, 1, 27-33.	REVIEW			

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
68. **Kanouse DE, Brook RH, Winkler JD, Kosecoff JK, Berry SH, Carter GM, Kahan JP, McCloskey L, Rogers WH, Winslow CM, Anderson GM, Brodsley L, Fink A, Meredith L. (1989)** Changing Medical Practice Through Technology Assessment: An Evaluation of the NIH Consensus Development Program. RAND publication no. R-3452-NIH/RC.	a	z	ai	b
69. **Karuza J. et al. (1995)** Enhancing physician adoption of practice guidelines: Dissemination of influenza vaccination using a small-group consensus process. Archives of Internal Medicine 155, 6, 625-632.	a, d	a	ai	a
70. **Kimura S, Pascala JT. (1997)** Pressure ulcers in adults: family physicians' knowledge, attitudes, practice preferences, and awareness of AHCPR guidelines. Journal of Family Practice 44, 4, 361-368.	a, d	a	ai, d	a, di, e, f, g
71. **Klabunde CN, O'Malley MS, Kaluzny AD. (1997)** Physicians' reactions to change in recommendations for mammography screening. American Journal of Preventive Medicine 13, 6; 432-438.	b	a	ai	a
72. **Kristeller JL, Hoerr RA. (1997)** Physician attitudes towards managing obesity: differences among six specialty groups. Preventive Medicine, 26, 4; 542-549.	e	z	[x]	b
73. **Kushner RF. (1995)** Barriers to providing nutrition counseling by physicians: a survey of primary care practitioners. Preventive Medicine 24, 6, 546-552.	e	a	ai	a
74. **Larue F, Colleau SM, Fontaine A, Brasseur L. (1995)** Oncologists and primary care physicians' attitudes towards pain control and morphine prescribing in France. Cancer 76, 11, 2375-2382.	e	a	ai	a
75. **Lekan-Rutledge D, Palmer MH, Belyea M. (1998)** In their own words: nursing assistants' perceptions of barriers to implementation of prompted voiding in long-term care. Gerentologist 38, 3, 370-378.	d	f	aiii	di

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
76. **Lennox NG, Diggens JN, Ugoni AM. (1997)** The general practice care of people with intellectual disability: barriers and solutions. Journal of intellectual disability Research 41, 5, 380-390.	e	a	a[x]	di
77. **Linde BJ. (1989)** The effectiveness of three interventions to increase research utilization among practicing nurses. PhD Thesis. The University of Michigan.	b	d	aiii	a
78. **Linn LS, Yager J, Leake B. (1988)** Psychiatrists' attitudes toward preventive intervention in routine clinical practice. Hospital and Community Psychiatry 39, 6, 637-642.	e	b	aiii	b, di, g
79. **Louis CZ, Taroni F, Yuen EJ, Umesato Y, Gonella JS. (1996)** Patterns of hospital care and physicians perspectives from an Italian, Japanese, and USA hospital. American Journal of Medical Quality 11, 3, 123-134.	d	z	ai	di
80. **Lomas J, Haynes RB. (1988)** A taxonomy and review of tested strategies for the application of clinical practice recommendations from 'official' to 'individual' clinical policy. American Journal of Preventive Medicine 4, 77-94.	REVIEW			
81. **Lomas J, Anderson GM, Domnick-Pierre K, Vayda E, Enkin MW, Hannah WJ. (1989)** Do practice guidelines guide practice? The effects of a consensus statement on the practice of physicians. New England. Journal of Medicine 321, 19, 1306-1311.	a	b	ai	a
82. **Lynn MR, Moore K. (1997)** Research utilization by nurse managers: current practices and future directions. Seminars for nurse managers 5, 4, 217-223.	b	d	a[x]	di
83. **Mansfield C. (1995)** Attitudes and behaviours towards clinical guidelines: the clinicians' perspective. Quality in Health Care 4, 250-255.	a	z	ao	a
84. **Margolis FJ, Chesney BK, Schork A. (1987)** Fluoride supplements. Changes in physicians' attitudes and practices following an intensive, multifaceted educational program. American Journal of Diseases of Children 141, 1, 72-76.	d	a	ai	a

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
85. **Martinez JL, Licea Serrato J de D, Jiminez R; Grimes RM. (1998)** HIV/AIDS practice patterns, knowledge, and educational needs among Hispanic clinicians in Texas, USA, and Nuevo Leon, Mexico. Revista Panamericana De Salud Publica 4, 1, 14-19.	e	a	[x]	a
86. **McKaig C, Hindi-Alexander M, Myers TR, Castiglia P. (1984)** Implementation of the school nurse practitioner role: barriers and facilitators. Journal of School Health 54, 1, 21-23.	e	e	bii	diii, div
87. **McColl A, Smith H, White P, Field J. (1998)** General practitioners' perceptions of the route to evidence-based medicine: a questionnaire survey. BMJ 316 (7128) :361-365.	b	a	ai	b, di, h
88. **Meah S, Luker KA, Cullum NA. (1996)** An exploration of midwives' attitudes to research and perceived barriers to research utilisation. Midwifery 12, 2, 73-84.	b	e	c	b, di
89. **Meyers DG, Steinle BT. (1997)** Awareness of consensus preventive medicine practice guidelines among primary care physicians. Am J Prev Med 13, 45-50.	a	a	ai	c
90. **Miller JR, Messenger SR. (1978)** Obstacles to applying nursing research findings. American Journal of Nursing 78, 632-634.	b	d	3a	d
91. **Miller KE, Miller MM, Single N. (1997)** Barriers to hospice care: family physicians' perceptions. Hospice Journal 12, 4, 29-41.	e	a	a[x]	di
92. **Mottur-Pilson C. (1995)** Internists' evaluation of guidelines: the IMCARE practice guidelines network. International Journal of Quality in Health Care 7, 31-37.	a	a	aiii	a
93. **Naylor CD, Hollenberg AA, Ugnat AM, Basinski A. (1990)** Coronary thrombolysis – clinical guidelines and public policy: results of an Ontario practitioner survey. Canadian Medical Association Journal 142, 1069-1076.	a	b	ai	a
94. **Newton J, Knight D, Woolhead G. (1996)** General practitioners and clinical guidelines: a survey of knowledge, use and beliefs. British Journal of General Practice 46, 410, 513-517.	b	a	a[x]	a, di

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
95. **Nilsson Kajermo K, Nordstrom G, Krusebrant A, Bjorvell H (1998)** Barriers to and facilitators of research utilization, as perceived by a group of registered nurses in Sweden. Journal of advanced Nursing 27, 4, 798-807.	e	d	ai	di
96. **Olesen F, Lauritzen T. (1997)** Do general practitioners want guidelines? Attitudes towards a county-based and national college based approach. Scandinavian Journal of Primary Health Care 15, 3, 141-145.	a	a	ai	b, c
97. **Olson JM. (1992)** Psychological barriers to behavior change. Canadian Family Physician, 38, 309-319.	REVIEW			
98. **Oxman AD, Thomson MA, Davis DA, Haynes R, Brian MD. (1995)** No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. Canadian Medical Association Journal 153, 10, 1423-1431.	REVIEW			
99. **Pardell H, Tresserras R, Armario P, Salto E. (1990)** Detection, treatment and control of hypertension in Spain. Journal of Human Hypertension 4, 4; 297-299.	d	a	ai	a, di
100. **Patton CB, Beaman M, Csar N, Lewinski C. (1996)** Nurses' attitudes and behaviours that promote breast-feeding. Journal of Human Lactation 12, 2, 111-115.	d	e	ai	diii, dv
101. **Pereles I, Russell ML. (1996)** Needs for CME in geriatrics. Part 2: Physician priorities and perceptions of community representatives. Canadian Family Physician 42, 632-640.	g	z	bii	f
102. **Pill R, Stott NC. (1988)** The Open University pack on coronary heart disease: reactions of Welsh general practitioners. Medical Education 22, 6, 543-549.	a	a	[x]	di
103. **Prescott K, Lloyd M, Douglas H-R, Haines A, Humphrey C, Rosenthal J, Watt I. (1997)** Promoting clinically effective practice: General practitioners' awareness of sources of research evidence. Family Practice 14, 320-323.	a, b	a	ai	c, i

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
104. **Renvoise EB, Hampshaw SM, Pinder JM, Ayres P. (1997)** What are hospitals doing about clinical guidelines? Quality in Health Care 4, 187-191.	a	b	aiii	b, di
105. **Robertson N, Baker R, Hearnshaw H. (1996)** Changing the clinical behaviour of doctors: a psychological framework. Quality in Health Care, 5, 51-54.	REVIEW			
106. **Rush BR, Powell LY, Crowe TG, Ellis K. (1995)** Early intervention for alcohol use: family physicians' motivations and perceived barriers. Canadian Medical Association Journal 152, 6, 863-869.	d	a	c, bii	di
107. **Rutledge DN, Ropka M, Greene PE, Nail L, Mooney KH (1998)** Barriers to research utilization for oncology staff nurses and nurse managers/clinical nurse specialists. Oncology Nursing Forum 25, 3, 497-506.	b	d	ai	di, g
108. **Scott-Connor CE, Christie DW, McGehee R, Anglin BL, Heatherington R. (1996)** Breast and cervical carcinoma screening services: physicians' perceptions of need. Journal of the Mississippi State Medical Association 37, 11, 809-815.	d	z	[x]	a
109. **Shi L, Samuels DL, Stoskopf CH, Baker SL. (1997)** Primary care physicians and barriers to providing persons with HIV/AIDS. Evaluation & the Health Professions 20, 164-187.	d	a	ai	a
110. **Shye D, Brown JB. (1995)** Primary care HMO clinicians' opinions about clinical practice guidelines. HMO Practice 9, 3, 111-115.	a	a	ai	b
111. **Siriwardena AN. (1995)** Clinical guidelines in primary care: a survey of general practitioners' attitudes and behaviour. British Journal of General Practice 45, 643-647.	a	a	aiii	a, g
112. **Soliman AS, Raouf AA, Chamberlain RM (1997)** Knowledge of, attitudes toward, and barriers to cancer control and screening among primary care physicians in Egypt: the need for postgraduate medical education. Journal of Cancer Education, 12, 2, 100-107.	e	a	[x]	a

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
113. **Soumerai SB, Avorn J. (1990)** Principals of educational outreach ('academic detailing') to improve clinical decision making. JAMA 263, 549-556.	REVIEW			
114. **Stanton WR, Balanda KP, Gillespie AM, Lowe JB. (1996)** Barriers to health promotion activities in public hospitals. Australian and New Zealand Public Health 20, 5, 500-504.	e	g	ai	di
115. **Stange KC, Kelly R, Chao J, Zyzanski SJ, Shank C, Jaen CR, Melnikow J, Flocke S. (1990)** Physician agreement with US Preventive Services Task Force Recommendations. Journal of Family Practice 34, 409-416.	a	a	ai	b, c, f, g
116. **Suarez-Varela Ubeda JF, Zuzunegui MV, Bimbela JL, Vilches A. (1997)** Attitude and structural barriers detected in primary care physicians towards treating patients with HIV/AIDS infection (In Spanish). Atencion Primaria 20, 9, 486-492.	e	a	[x]	a
117. **Swinburn BA, Walter LG, Arrol B, Tilyard MW, Russel DG. (1997)** Green prescriptions: attitudes and perceptions of general practitioners towards prescribing exercise. British Journal of General Practice 47, 422, 567-569.	d	a	c	di
118. **Togno JM, Lundin R, Buckley P, Hovel J. (1996)** Rural health and IT&T in Australia – the results of qualitative and quantitative surveys of the needs, perceptions and expectations of rural and remote health professionals. Journal of Telemedicine and Telecare 2, 1, 104-105.	e	z	aiii	di
119. **Tseng FY. (1996)** Quality of medical care for non-insulin-dependent diabetic patients in a community teaching hospital in Taiwan. Journal of the Formosan Medical Association 95, 5, 414-416.	a	b	[x]	di
120. **Tunis SR, Hayward RSA, Wilson MC, Rubin HR, Bass EB, Johnston M, Steinberg EP. (1994)** Internists' attitudes about clinical practice guidelines. Annals of Internal Medicine 120, 11, 956-963.	a	a	ai	a, g
121. **Veit FC, Sanci LA, Coffey CM, Young DY, Bowes G. (1996)** Barriers to effective primary health care for adolescents. Medical Journal of Australia 165, 3, 131-133.	d	a	ai	di, h

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
122. **Ventura MR, Feldman MJ, Crosby, F. (1989)** Nurse practitioner perceptions of facilitators and constraints to practice. Military Medicine 154, 6, 303-311.	e	e	ai	di
123. **Wallace KG, Reed BA, Pasero C, Olsson GL. (1995)** Staff nurses' perceptions of barriers to effective pain management. Journal of Pain and Symptom Management 10, 3, 204-213.	d	d	ai	diii, dvi
124. **Walsh M. (1997a)** Perceptions of barriers to implementing research. Nursing Standard 11, 19, 34-37.	b	e	ai	di
125. **Walsh M (1997b)** How nurses perceive barriers to research implementation. Nursing Standard 11, 29, 34-39.	b	e	ai	di
126. **Watkins C, Harvey I, Langley C, Gray S, Faulkner A. (1999)** General practitioners' use of guidelines in the consultation and their attitudes towards them. British Journal of General Practice, 49, 11-15.	a	a	aiii	b, d
127. **Weingarten S, Stone E, Hayward R, Tunis S, Pelter M, Huang H, Kristopaitis R. (1995)** The adoption of preventive care practice guidelines by primary care physicians: do actions match intentions? Journal of General Internal Medicine 10, 3, 138-144.	a	a	ai, e	a
128. **Weingarten S, Riedlinger M, Bolton LB, Miles P, Ault M. (1989)** Barriers to influenza vaccine acceptance. A survey of physicians and nurses. American Journal of Infection Control, 17, 4, 202-207.	d	z	[x]	di
129. **Wensing M, Hulscher M, Laurant M, Grol R. [1999]** Identifying barriers and facilitators for changing professional practice. Nijmegen: Centre for Quality of Care Research (WOK).	REVIEW			
130. **Wolff M, Bower DJ, Marbella AM, Casanova JE (1998)** US family physicians' experiences with practice guidelines. Family Medicine 30, 2, 117-21.	a	a	ai	a
131. **Wright J, Wilkinson J. (1996)** General practitioners' attitudes to variations in referral rates and how they could be managed. Family Practice 13, 3, 259-263	e	a	aii, bii	di

Source	*1. Focus*	*2. Targets*	*3. Methods*	*4. Object*
132. **Zucker AR, Meadow WL. (1995)** Pediatric critical care physicians' attitudes about guidelines for the use of ribavirin in critically ill children with respiratory syncytial virus pneumonia. Critical Care Medicine 23, 4, 767-772.	a	b	aiii	b, di

Chapter 10

Epilogue

Marjukka Mäkelä and Thorkil Thorsen

The provision of health care services is a complex social process. In each encounter between patient and professional, there are multiple issues that influence the decisions made and the outcomes produced. Naive assumptions about the transference of scientific evidence into work practices do not hold: high-quality scientific evidence is not always used as a natural ingredient in health care decisions. Even the ready availability of high-quality guidelines does not necessarily improve this situation. The path from available scientific knowledge to the application of this knowledge in patient care is winding and stony.

In early 1995, we saw a gap rather than a path. A small group of guideline researchers was called together by Finn Børlum Kristensen at the Danish Hospital Institute. All were planning to start a study on guideline implementation in the near future. Likewise all agreed that important issues in this work could best be studied by joining resources. The bits of understanding gained from our various perspectives would support each other, creating a chance to ask more complicated questions. At that point we were clearly aiming at using a common vocabulary and applying some common instruments in our respective study projects.

After three years of active and mostly very satisfying collaboration we have some new pieces of knowledge in place. First, we know much more about designing and analysing guidelines implementation studies. Tools for using cluster design and analysis have been sharpened and tested using the data from CPP studies. The application of economic evaluation specifically to guideline implementation studies has been carefully considered; the conclusion is that the methods are appropriate and should be included in study designs. There is also a clear need for richer description of the complex processes we study; the next genera-

tion of guideline implementation studies should be able to consider qualitative approaches – unlike CPP at the outset!

A second tangible result of the collaboration is the scientific framework that can be used in guideline implementation studies, either just for planning the implementation process itself or also for designing a study. The framework is based on original work by EPOC, which aimed at retrospective analysis of published studies; we have elaborated on this using data from CPP partner projects. To support study planning instead of analysis, and to incorporate the new insights from research, the framework has been enlarged with a more detailed list of guidelines, interventions and intervening factors. Again, future researchers now are in a better position to choose the setting, persons and guidelines they work with than we were when our projects were designed.

Thirdly, seminal theoretical work on the description of interventions and the identification of barriers to and facilitators of guideline implementation has been done in the Netherlands. These results enrich the main framework. Likewise, during the Concerted Action a bibliography was compiled of studies investigating health professionals' perceptions of factors that may interfere with guideline implementation and use.

A systematic review of studies looking at guideline implementation has not been done yet. We planned to do one with our own results. Too optimistic! Several of the projects that originally were included in CPP are still in the process of being executed; not all of them were even funded in the end. We think, however, that our work has made a solid systematic review possible, and will keep our eyes open for the right time to start one.

Another of our originally planned project deliverables, the common survey instrument looking at barriers to and facilitators of guideline implementation, has not been as widely applied as we scheduled. Developed using previous experience and tested on two sets of empirical material, the CPP-14 questionnaire is available in English, Finnish, Spanish and Danish, but not in all the languages used in CPP. Neither have all projects used it, although this was an initial goal. The questionnaire needs to be validated for countries and settings other than the one it was tested in.

The most useful and hopefully most permanent result of CPP is the creation of a European network of guideline enthusiasts. The people and groups doing research and development in the area now have good knowledge of each other's professional strengths. We have been able to share work fruitfully and question our results honestly. Most importantly, we have gained insight into questions that none of us would have

been able to ask within the limits of one's own country and health care system. 'Added value' for us has changed from a BIOMED mantra to a very tangible fact.

Some of the many lessons learnt during our Concerted Action, we believe, grew from the very idea of multinational European collaboration. There is a clear need to systematise approaches and results. While it is very true that cultural contexts and contextual issues are important in health services research, many problems are nevertheless generic across very different guideline implementation studies. Teams with members that have multidisciplinary breadth as well as depth in their own fields seem optimal for this branch of health services research. Cultural variety is a true richness, when we first dare pass the threshold from fear of the unknown and protection of our familiar ways to honest and fearless exchange – and when we trust our partners enough to laugh at ourselves and each other. We need to continue teasing out generic explanations and hopefully also solutions to guideline implementation.

The path from available scientific knowledge to applying this knowledge in patient care is still winding and stony. The theoretical and practical work done within CPP has cleared some stones away, and provided a few signposts. We are working on numerous new questions with a better map of the terrain than we used to have, and hope that others will join our network in the quest for more scientifically based health care.

Appendix

Partner projects

The following projects were included in the CPP Concerted Action and are referred to in the previous chapters:

LysAMI: Optimising the treatment of myocardial infarction. Research into implementation of clinical guidelines

Denmark 1

Background

Every year approximately 16,000 Danes suffering from myocardial infarction (MI) are discharged from one of more than 60 hospital departments that admit patients with MI. MI causes 20% of all deaths. Half of these deaths occur during hospitalisation. Extensive clinical trials have shown that treatment with acetylsalicylic acid and, when indicated, earliest possible thrombolytic treatment has increased survival in MI-patients. Other procedures of risk assessment and subsequent interventions have been shown to increase survival. This highlights the importance of a thorough clinical risk assessment in MI patients. It is not well described to which extent this international knowledge of most effective treatment is reflected in day-to-day clinical practice in Denmark.

Aim

An improvement of clinical guideline implementation, thus optimising the treatment of MI.

Study design

In a rigorous health services research design, this study was to address the area of quality assurance and improvement through guideline implementation. A prospective interrupted time series design was to be employed to test two methods of guideline implementation in hospital

care against each other: statistical feed-back on performance indicators vs. statistical feed-back plus a centrally initiated and facilitated local multidisciplinary quality development work.

The intervention was also to be studied in an organisational/sociological part study. The goal here was to study a) what happens when a hospital department systematically records its treatment of patients and receives a feedback, and b) how is the clinical guideline received, perceived and used in the departments. A special focus was to be on barriers and facilitators to the implementation of the guideline.

The study has documented the process of care of in-patients admitted with MI symptoms during a pre-intervention period of 8 months. Thus, the MI-treatment profile of each department included in the study has been recorded. During and after the intervention period registration of the MI-treatment activities in all departments were to continue unchanged.

The *end point measure* used for evaluating the impact of the implementation strategy was primarily the health professionals' compliance with the clinical guideline on well-established performance indicators. For example, it was to be expected that a successful implementation of the guideline would result in an increased thrombolysis treatment rate and a quicker process of care in the group of patients who met the treatment's indications. There is a well-documented causal relation between specific treatments such as thrombolysis, beta-blockade, and cholesterol lowering and patient outcomes such as mortality. The use of process measures rather than patient outcomes as measures of effect calls for much less cases to be included. Thus, process of care measures could substitute patient outcomes as primary effect measures in this study. This means that well documented *therapy* (such as aspirin and thrombolysis) and – with a more indirect causal relationship – pre-discharge and ambulatory *diagnostic procedures* necessary for risk assessment prior to decision on therapy (such as exercise test, echocardiography, hypertension) would serve as measures of effect of the intervention.

The organisational/sociological part study was to be based on interviews and observation.

Population

Patients admitted to hospital with chest pain.

Progress

In January 1997, 26 hospital departments had been enrolled in the study. Each department had established an organisation for data collection in collaboration with the project group and appointed a contact person re-

sponsible for the continuous recording of their treatment of patients with MI-symptoms. In the following 9 months the departments' registration practices were being checked and improved upon before what was supposed to be the beginning of the pre intervention period. Not all departments managed to obtain a quality of data that was acceptable for the study but 25 departments managed to start the collection and most at a reasonable quality.

Nevertheless, as it had not been possible to obtain the necessary funding, the intervention was not started as planned, and in October 1997 the continuous recording was officially stopped in all departments due to lack of funding. A few departments chose to continue on their own but it was no longer possible to support the recording at DSI. At that time 12,500 registration forms had been scanned into a database, and a feedback on their recording was sent to all departments. Since then a further 3,000 forms have been scanned.

Further information:

DSI•Danish Institute for Health Services Research and Development, P.O. Box 2595, DK 2100 Copenhagen Ø.
Phone: +45 35 29 84 00. Fax: +45 35 29 84 99.

A randomised controlled trial of computer-assisted anticoagulant therapy

Denmark 2

Background

Outpatient anticoagulant therapy with peroral vitamin-K antagonists is often of poor quality. One of the reasons seems to be that the clinicians do not comply with guidelines and other evidence. In an attempt to improve the quality, a computer-system to assist the clinician during the patient's periodic control visits is being developed and implemented. The computer-system consists of:

- A patient record interactively connected to a clinical guideline and a dose calculation program,
- An information database easily accessible from the patient record, and
- A regional (and eventually national) clinical database for the clinical intervention and its results.

Aims

The primary objective of this study is to evaluate the change of professional behaviour of the clinicians, and the resulting quality improvement of the therapy. The hypotheses are:

1. That the clinicians will use the computer-system during the encounter,
2. That the reminders and information received will change the professional behaviour, and
3. That this will improve the clinical outcome.

Study design

The evaluation is a randomised controlled study. The intervention group receiving the computer-system consists of about 25 general practices and 6 outpatient clinics in Funen County. Another 25 general practices in Funen County and a number of outpatient clinics outside Funen constitute the control group. Referring to the above hypotheses, the parameters measured are among others:

1. Frequency of the use of individual parts of the computer-system,
2. Frequency of control visits, duration of treatment, and frequency of patient information provided,
3. Relative duration of which the patient's INR is within the therapeutic range.

Primarily data are collected by computer data logging, and from laboratory and hospital information systems. This is followed up by post intervention questionnaires and interviews.

Population / Setting
General practices and outpatient clinics mainly in the Funen County, Denmark.

Progress
Development and implementation of the computer tools was delayed 6 - 9 month due to lack of human resources. Whereas development and implementation of computer-assisted decision support (clinical guideline, database, etc.) is part of the PRESTIGE Project, funded by the European Commission, national funds and partners, funding of the evaluation of the effect of the of computer-assisted decision support (i.e. this project) is not secured, and can only to a very limited extent be covered by the project partners.

Further information:
DSI•Danish Institute for Health Services Research and Development, P.O. Box 2595, DK 2100 Copenhagen Ø.
Phone: +45 35 29 84 00. Fax: +45 35 29 84 99.

Quality management methods in health centres (Laadukas terveyskeskus, LATE)

Finland 1

Background

Quality improvement (QI) methods have been increasingly applied in health care, with little data on their effectiveness. Finnish primary care units have been interested in remodelling their work toward truly multiprofessional teams; the implementation of guidelines is a suitable topic for quality improvement in such changes. This project combines the practical application of guidelines with an educational program aiming at facilitating multiprofessional quality improvement in teams.

Aims

The study questions are:

1. Will the application of QI methods in primary care lead to
   a) improved health outcomes in patients with diabetes and hypertension and/or
   b) improved care processes in patients with tonsillitis and sinusitis?
2. What are the costs of implementing this type of team education?
3. What are the attitudes of the health centre personnel toward guidelines before and after the implementation?

Study design

In this randomised trial the effectiveness of teaching QI methods to multiprofessional teams in primary health care is evaluated. Six health centres (HCs) started in 1996 and six in 1997. Each HC receives a training program extending over 18 months, including four one-day sessions locally in the health centres and five common training sessions for five persons chosen by the HCs, two quality managers, a physician and a nurse, and three quality project leaders. National guidelines are used as the basic source of information in all QA projects.

In the trial, each HC receives two randomised topics for QA. One of these is a chronic health problem (hypertension or diabetes) and the other an acute problem (sinusitis or tonsillitis). In addition, the HCs may choose one extra QA topic according to their own preferences. All HCs collect retrospective data on the actual treatment of 200 patients (or all available patients, if the total n<200) with each of these health problems yearly in 1996 to 1999. The hard outcomes are the percentages of patients with their latest blood pressure or blood sugar level in good control, or patients who have received a treatment for their acute respi-

ratory infection correctly according to the results of a diagnostic investigation.

Progress

The project is nearing completion. The post-intervention measurements are still being collected; they have taken somewhat longer than expected. Plans for the analysis have been made and funding for the personnel and statistics experts necessary for the final analysis are being sought for. The data are expected to be available by the end of 1999.

All training sessions for the HCs that started in 1997 are completed, with the final evaluation session in June 1999. All HCs have completed the data collection for 1997 and all but two for 1998. Baseline data (1996) for the all HCs have been analysed and a conference report on these presented. Preliminary reports on first-year data, with trends for target problems, have been written. Attitude questionnaires have been distributed to all HCs; baseline data on attitudes has been analysed. Three yearly project reports have been published and distributed. A basic scientific report for the project has been started: the literature review for it is completed, and basic tables for reporting the data have been planned.

Other

The training program was tested in six HCs in 1995-97 and small modifications to the program were made after that. Health centres receive feedback on their progress and the comparative figures from other HCs during the project; however, these data are available so late that its effect on the improvement process is marginal.

The training program will continue in a slightly modified format, with six new HCs starting in the autumn of 1999. Data will not be collected and HCs will be free to choose their project topics.

Further information

Marjukka Mäkelä, STAKES/PALA, Broholmsgatan 18, PB 220, 00531 Helsinki, Finland.
Tel: +358 9 3967 2290. Fax: +358 9 3967 2227.
e-mail: Marjukka@stakes.fi

Primary care guidelines in electronic vs. textbook format

Finland 2

Background

Electronic guidelines have been used in Finnish primary care for seven years. As primary health centres (HCs) become increasingly computerised, it is essential to evaluate the effect of readily available guidelines on the quality of general practitioners' work. An exceptional possibility to compare a collection of two formats (printed and electronic) of a collection of over 1000 guidelines for primary care, and willingness to evaluate in detail how their recommendations are being followed, prompted this randomised study.

Aims

The study questions are:

1. How well do primary care physicians follow the Physician's Desk Guideline in their daily work, as evaluated from patient records from cases where guidelines are consulted vs. cases were guidelines were not used?
2. Is there a difference between using electronic or printed (book) versions of the guidelines collection?
3. What are the attitudes of the physicians toward guidelines before and after the study?

Study design

In this study, newly licensed physicians are randomised to use either a textbook or an electronic version of existing guidelines for primary care. These guidelines have been produced since 1989 and are updated three times a year (database version) or every two years (printed version). The young physicians will receive either the book or a portable computer with the guidelines for use during one month in a health centre (their second or later month in primary care work after licensing). For each consultation where they use the guidelines they will fill in a short questionnaire, and a copy of the patient record for this consultation is included. The patient record for the previous consultation will be collected for control material.

The data will be analysed for the frequency of guideline use, problems guidelines were used for, sources of guidelines, whether the information provided in the guideline was sufficient for decision, and whether the physician acted according to the guideline. The latter two aspects are analysed by the research team from the patient records. The expected minimum number of physicians to be recruited is 166, and the minimum number of consultations to be analysed is 1112.

Progress

A pilot study in June 1997 has been analysed by four researchers independently to test the validity of the appraisal instruments. The recruitment for the project started in January 1998 and will end in August 1999; 136 physicians have been randomised for the study and 96 have completed the data collection. The number of available consultations exceeded the estimated number, so that a selection of conditions to be analysed became possible. The evaluation scales for the final analysis are being created for the 30 most common diagnoses. Index and control visits will be blinded for those doing the analysis. The guideline attitudes questionnaire has been distributed to those randomised.

Other

The guidelines are being translated into English for a demonstration version, and the CD-ROM containing the guidelines in English will be available for those interested in 1999.

Further information

Duodecim (The Finnish Medical Society) and Stakes (National Research and Development Centre for Welfare and Health). Marjukka Mäkelä, Stakes, Broholmsgatan 18, PB 220, 00531 Helsinki, Finland.
Tel: +358 9 3967 2290. Fax: +358 9 3967 2227.
e-mail: Marjukka@stakes.fi

Audit and feedback vs. academic detailing as a strategy to implement practice guidelines for breast cancer surgery in Italy

Italy

Background

Studies carried out in the early 90′ showed that the rates of use of breast conserving surgery were variable and particularly low in some geographic areas. In particular, low rates of use of this procedure were found in Friuli Venezia Giulia, a Northern Italian Region with approximately 1,500,000 residents. This finding could not be attributed to the lack of radiotherapy facilities, rather to clinicians′ attitudes favouring the utilisation of more radical interventions. This indicated that there was much room for improvement in this clinical area through the development and implementation of practice guidelines promoting the use of conserving surgery in the appropriate patient with breast cancer. These guidelines were developed in 1995 by a multidisciplinary panel of regional experts dealing with management of breast cancer, using the RAND appropriateness methodology. Through this approach it was possible to reach a consensus on the clinical indications in which the use of breast conserving surgical procedure was deemed appropriate.

Aims

To assess the feasibility of setting up routine and simple mechanisms through which guidelines implementation could be organised by the regional Health Authority and to assess the comparative effectiveness of two different strategies for implementing guidelines aimed at increasing the use of breast conserving surgery (BCS).

Study design

Controlled before after study, involving 6 hospitals allocated to receive the guidelines either through audit and feedback (2 hospitals) or academic detailing (2 hospitals), and the remaining 12 hospitals of the region acting as control group. Audit and feedback: after a local presentation of the guidelines by the chief of the regional health authority, hospitals in this group have been audited every four months for 12 months, in order to monitor their use of breast conserving surgery. A report of their performance in terms of total number of breast cancer patients treated during the index period, number of women eligible for BCS according to the guidelines, and number of those who actually received BCS, was regularly provided to the chairperson of the surgical department. This material was then used at regular department meetings.

Academic detailing: physicians treating breast cancer patients at each hospital allocated to this arm of the study were invited to a meeting where an expert in the field championed the content of the guideline, discussed their method of development, content and potential problems in the local application. During this meeting the chief of the regional health authority and the expert presented and discussed data on the local utilisation of BCS during the previous year, and stressed the importance of the guideline adoption as a tool to improve quality of care. This meeting was offered only once, before starting the data collection.

The proportion of breast cancer patients treated with conservative surgery was the main outcome measure, estimated from data collected through the regional administrative database. Data from 1992 to February 1998 were available for the analysis.

According to the baseline rate of use of BCS, it was estimated that 360 surgical interventions had to be observed in order to detect a 15% difference between the two intervention arms (assuming 1-beta=80% and alfa=5%); considering the two implementation strategies altogether, 141 patients per group had to be recruited to detect a 15% increase in the rate of use of BCS compared with the control group.

Clinicians and staff of the surgical wards of the regional hospitals were the target population of the study.

Timetable

The two implementation strategies were launched in September 1996 and ended in November 1997. More specifically, written report in the audit and feedback were sent three times (Feb. 97, covering the period Sept-Dec 96, July 97, with data covering the period Jan-April 97; Nov 97 covering the period May-August 97). Final analysis was carried out in March 1998.

Progress

During the first year the study was carried out successfully though patient accrual was slower than expected. It was only at the end of 1997 that due to the turnover of personnel at the regional Health Authority difficulties started to emerge. As it is reported in detail below data collection ended at the end of 1997 and it proved impossible to rescue it afterwards also due to the lack of external funding.

Overall we examined rates of use of BCS in the control and experimental hospitals from 1992 to 1995, before the launch of the interventions. During that baseline period, in control hospitals rates went from 44% in 1992 to 40% in 1997. Corresponding rates went from 43% to 44%, and from 35% to 41%, in the audit and feedback and in the aca-

demic detailing group, respectively. Since the inception of the interventions, rates went from 40% in 1996 to 42% in 1997 in the control group; in the audit and feedback group corresponding figures were 34% in 1996 and 46% in 1997 while in the academic detailing group were 47% and 48%.

Due to the small sample size none of these differences was statistically significant. Although the study protocol foresaw a monitoring of BCS rates until 1999, data collection was interrupted prematurely at the end of 1997 due to organisational reasons at the Regional Health Authority.

Dissemination (publications – issued/planned)
We considered the study not successfully completed and therefore no formal publication is planned. An internal report has been produced and is available in Italian.

Further information
Alessandro Liberati, Istituto Mario Negri, Via Eritrea 62, 20157 Milan, Italy.
Tel: +39 2 390 145 16, Fax: +39 2 332 002 31.
e-mail: LIBERATI@IRFMN.MNEGRI.IT,

Tailored interventions in primary care

Norway

Background

Sore throat and urinary tract infections (UTIs) are common problems in general practice. Treatment and diagnosis of these conditions has been controversial, and there have been large variations in practice.

Aim

To evaluate the effect of tailored interventions on the implementation of evidence-based guidelines for the diagnosis and treatment of sore throat and UTI.

Methods and study design

Evidence-based guidelines for the diagnosis and treatment of sore throat and UTI were systematically developed. 140 practices will be randomised to an intervention to support implementation of the guidelines for either sore throat or UTI in a balanced block design.

Interventions

"Tailored interventions" that address specific barriers to appropriate practice for each clinical problem will be used. Focus groups and interviews among patients, general practitioners (GPs) and GP assistants, and a pilot study will be used to identify the barriers. Likely components of the tailored interventions include educational material for the practice focusing on daily routines, computer generated patient information, computerised decision support and reminders, use of opinion leaders and individual feedback.

Participants/setting

140 GP practices in Norway with 20 patients per practice will be included. The sample size is calculated to achieve 80 % power to detect a 15 % absolute difference in practice for the primary analyses. We assumed a "worst case" baseline of 50 % appropriate practice, an intra-class correlation coefficient of 0.3 and a drop-out of up to 20% of practices.

Measurement of the process of care, patient outcomes, economic evaluation

Symptoms and signs, use of laboratory tests, and prescription of antibiotics for both clinical problems will be registered electronically in all participating practices during three time periods of two months: before the interventions, during active implementation of the tailored inter-

ventions and one year after the active intervention phase. An add-on computer program that is minimally obtrusive will be used to collect data, prompt GPs and GP assistants, provide patient information, and facilitate audit and feedback. Principle outcomes are the rate of test use, rate of use of antibiotics, and rate of telephone consultations. The following patient outcomes will be registered using a mailed questionnaire: symptoms, side effects of treatment, use of patient information and degree of patient satisfaction.

Plan of analysis
Intention to treat analysis using generalised linear modelling or multilevel modelling to account for clustering of data will be used. Exploratory analyses will be undertaken to investigate whether identified barriers and facilitators were effect modifiers.

Progress
Guidelines for sore throat and UTI and the add-on computer program have been developed. We have started the pilot project and identification of barriers and facilitators.

Further information
Signe Flottorp, Health Services Research Unit, National Institute of Public Health, PO Box 4404 Torshov, N-0403 Oslo, Norway.
Tel +47 22 04 23 66, Fax +47 22 04 25 95,
e-mail: signe.flottorp@labmed.uio.no

Cardiovascular Risk reduction in Primary carE (CARPE)

The Netherlands

Background

Health problem: Prevention of cardiovascular disease in general practice among patients with cardiovascular diseases or risk factors (hypertension, cholesterol, diabetes mellitus II, peripheral arterial disease, angina pectoris, heart failure and TIA). Cardiovascular diseases are among the most important health problems in general practice: about 10% of all chronic problems are of cardiovascular origin.
Effectiveness: Aiming preventive activities at patients with established disease or at risk, is a cost-effective approach, in line with current literature and discussions.
Evidence of current inappropriate practice: Research in different countries documents discrepancies between evidence based preventive guidelines and objective measures of what is done in practice.

Aim

To evaluate the effects of outreach visits by trained nurses to implement evidence based performance guidelines and guidelines on organising cardiovascular preventive services.

Study design

Design
Randomised controlled trial.

Guidelines
The Dutch College of General Practitioners has developed guidelines for the management of cardiovascular diseases and risk factors in their 'standards' on hypertension, cholesterol, diabetes mellitus II, peripheral arterial disease, angina pectoris, heart failure and TIA. These 'standards' provide specific performance and organisational guidelines. Moreover, guidelines to organise services have been developed and tested in a previous project on improving cardiovascular preventive care. Together this package of guidelines forms a suitable point of reference for improving patient care in this very important field.

Intervention
Multifaceted intervention, combining the provision of information, feedback, face-to-face education, educational materials and practical tools. The intervention is carried out by trained nurse outreach visitors.

Participants and setting
124 general practices (general practitioners and practice nurses) in the eastern and western part of the Netherlands.

Measurement of the process of care, health outcomes, economic evaluation
Measurements are performed at pre-test and post-test (i.e. after 21 months of intervention). Data on the organisation of services are collected by means of observation of routine procedures and questionnaires for the general practitioner (GP) and the practice nurse. Principal outcome is proportion of general practices adhering to the guidelines. Concerning the providers' performance, adherence to the guidelines is measured by self-recording forms, which have to be completed immediately after a consultation (during a period of 2 months). Principal outcome is proportion of general practitioners adhering to the guidelines. Patients receive a questionnaire at baseline and after 21 months (principal outcomes: reported risk profile, quality of life, satisfaction with the care provided). At post-test, patient records are searched to assess the number of recorded risk factors. During the intervention, a continuous process evaluation takes place to measure the input of outreach visitors and practice members (economic evaluation).

Barriers and facilitators
At post-test, information on barriers and facilitators is collected.

Monitoring
Implementation of the intervention activities is monitored.

Sample size
124 general practices are randomised to an intervention (62 practices) and a control group (62 practices). A total of 199 general practitioners, 250 practice nurses and 4000 patients participate in the study.

Plan of analysis
Intention to treat analysis. Organisational guidelines: sign test or Mann-Whitney *U* test, to test difference of proportions. Performance guidelines: if possible, multi-level analysis to account for clustering of data. Descriptive analysis of barriers and facilitators and costs of the intervention.

Progress
All 124 practices participate in the study. Postmeasurements will take

place until July 1999. Data entering will end by August 1999 and data analyses will start subsequently.

Further information:
Marlies EJL Hulscher, Centre for Quality of Care Research (WOK),
University of Nijmegen, PO Box 9101, 6500 HB Nijmegen.
Tel: +31 243613129, Fax: +31 243540166,
e-mail: M.Hulscher@HSV.KUN.NL.

Thrombo-Embolism Prevention Evaluation Study (TEMPEST)

Aberdeen, UK

Background

Health problem: Deep vein thrombosis of the lower limb (DVT) and its acute complication, pulmonary embolism (PE) are major causes of death and disability in developed countries.
Effectiveness of clinical intervention: Routine prophylaxis in moderate or high-risk groups is the most efficient management option.
Evidence of current inappropriate practice: Audits in UK and US hospital settings suggested that high-risk patients were not receiving prophylaxis.

Aims

To evaluate the effectiveness and efficiency of three different dissemination and implementation strategies for evidence based guidelines for the prevention of deep vein thrombosis.

Study design

Design
$2 \times 2 \times 2$ factorial randomised controlled trial. One-year feasibility study undertaken.
Intervention: Guideline: national guidelines for prophylaxis of deep vein thromboembolism based on systematic review, developed by multidisciplinary group and evidence-linked recommendations.

Interventions
i. Educational outreach visits to ward staff;
ii. Reminder systems at the time of consultation;
iii. Role substitution involving nursing or pharmacy staff assessing patients' risk status.

Participants
Health care workers in hospital settings – senior and junior doctors, nurses, pharmacists.

Settings
Up to 66 clinical directorates in hospitals in Scotland and northern England.

Measurement of the process of care, health outcomes, economic evaluation
Data collected by case note review from pre and post intervention cohort of medium and high-risk patients identified from routine discharge data. Principle outcome is proportion of patients receiving thromboprophylaxis. Health outcomes not measured. Additional information for economic evaluation from case note review, published data, standard unit costs and expert opinion.

Barriers and facilitators
Prospective assessment of barriers during the feasibility study period. Monitoring of implementation strategies.

Sample size
30 patients per directorate from a minimum of 66 directorates (total number of patients 1,980) needed to achieve 80% power to detect a 20% change in prophylaxis rates (from 40% to 60%) using a 5% significance level and assuming an intra-class correlation of 0.3.

Plan of analysis
Intention to treat analysis using generalised linear modelling or multilevel modelling to account for clustering of data. Exploratory analysis of whether identified barriers and facilitators are effect modifiers. Cost benefit analysis using a balance sheet approach for economic evaluation.

Progress
Feasibility study completed involving recruitment of over 17 hospitals, widespread interest in trial expressed by other hospitals; methods of data collection proven; survey of barriers using 'Theory of Planned Behaviour' framework completed. Baseline prophylaxis rates higher than expected (>90% in surgical directorates, 75% in medical directorates), and thus, a move to the main trial was not regarded feasible.

Further information
Jeremy Grimshaw, Health Services Research Unit, University of Aberdeen, Medical School, Foresterhill, Aberdeen AB25 9ZD.
Tel: +44 01224 663123 ext. 54105, Fax: +44 (0)1224 663087,
e-mail: j.m.grimshaw@abdn.ac.uk

COGENT: An evaluation of computerised guidelines for the management of two chronic conditions in UK primary care

Newcastle, UK

Background

Clinical guidelines are being increasingly developed to promote quality of care and to reduce inappropriate variation in clinical practice. However, there is uncertainty about both their effectiveness and how best to introduce them into practice. A recent review suggested the most powerful implementation strategy involved patient-specific prompts at the time of consultation. Advances in primary care computing now allow the generation of interactive prompts based upon relatively complex guidelines to be presented on the screen during consultations. Ischaemic heart disease and asthma in adults are common chronic illnesses associated with high levels of morbidity and mortality and cared for predominantly in primary care.

Aim

The study will evaluate the impact of the development and implementation of interactive computer-based guidelines on the process and outcome of care for adult patients with ischaemic heart disease or asthma cared for in UK primary care.

Description of the study

Design

A before and after cluster randomised controlled trial using a 2×2 balanced incomplete block design

Guideline

National guidelines created from systematic reviews in a group comprising experts in primary care, secondary care, patients, public health physicians and economists. The group method was informal. The end users participated. The guideline has been updated and was not pretested or tailored locally.

Intervention

The guidelines were formatted into a decision support system to deliver patient specific prompts at the time of consultation.

Participants and settings

General practitioners and practice nurses in 64 UK general practices; 40 patients with each condition per practice.

Measurement of the process of care, health outcomes, economic evaluation
Data on the process and outcome of care will be gathered from patients records. Data on patient health outcomes will be gathered by self-completion postal questionnaire using both generic and condition specific instruments. Data on the costs of implementing the guidelines and their impact on the costs of care will be collected.

Barriers and facilitators
Data on potential barriers and facilitators was gathered within a previous project; an embedded case study will examine these issues within the trial.

Sample size
The study will have 80% power to detect a 10% difference in compliance with the guideline recommendations at a 5% level of significance

Plan of analysis
The analysis will use multi level modelling focussing on the practice as the unit of analysis. It will estimate compliance with the guidelines before and after the implementation.

Further information
Martin Eccles, Professor of Clinical Effectiveness, Centre for Health Services Research, University of Newcastle upon Tyne, 21 Claremont Place, Newcastle upon Tyne NE2 4AA, UK.
Tel: +44 191 222 6780. Fax: +44 191 222 6043.
e-mail: Martin.Eccles@ncl.ac.uk

CAMBIE: Changing physicians' behaviour in the management of unstable angina through a clinical practice guideline

Seville, Spain

Background

Unstable angina (UA) is a frequent and severe condition in which medical skills are particularly crucial for the patient. The clinical management of UA is described as a decision cascade with different steps such as early recognition, appraisal of the severity and hemodynamic factors, better approach for risk assessment and, finally, the decision of treatment (medical or interventionist). Furthermore, this process can be attended in different settings and medical specialities, including emergency services, intensive care units or inpatients wards. There is good evidence that an appropriate management can improve patient outcomes, including fatal and non-fatal complications.

A previous study in Seville hospitals showed a high rate of inappropriateness in the use of coronary angiography, according to criteria set by an expert panel, and a high variability among hospitals and medical specialities in the use of different procedures.

Aim

To evaluate the effectiveness of a strategy to develop and implement an evidence-based guideline for the management of unstable angina.

Description of the study

Design

A before and after stratified cluster controlled trial.

Guideline

Locally developed by a multidisciplinary group, including intended users. A clinical algorithm of the process of care (from the initial suspect to treatment) provided the basis for a systematic review of each clinical decision, complemented with the results of a previous panel of experts, using the RAND method.

Intervention

Academic detailing in each practice unit, adaptation to local circumstances, and feedback on former practice. Control group: passive diffusion.

Participants and settings

All (144) cardiologist and internal medicine physicians in ten practice

units of the three university hospitals of Seville (Spain) are randomised to intervention or control group.

Measurement of the process of care and outcomes
Data collected from patients records before and after the intervention and identified from routine discharge data. Main outcomes: compliance with recommendations and appropriateness of the management.

Barriers and facilitators
Pre-intervention attitudes surveyed by the CPP-18 instrument. Monitoring of implementation strategies.

Sample size
436 patients in each group (total: 872) needed to achieve 80 % power to detect a 15 % change in rates of appropriateness (60 to 75 %), with a 5 % significance level and assuming an intra-class correlation of 0.05.

Plan of analysis
Focused on compliance with guideline in each episode of care, before and after the implementation. Intention to treat principle will be applied. Analysis using multi-level modelling to account for clustering and covariates effect.

Progress
The guideline have been successfully implemented and endorsed by Spanish professional societies. Methods of data collection validated. Pre-intervention data already collected from patients records and post-intervention in progress. Spanish version of AGS-14 translated and validated. Survey finished with a response rate of 76.4 %, analysis in progress. End of data collection and analysis planned for September 1999.

Further information
Eduardo Briones, Agencia de Evaluación de Tecnologías Sanitarias de Andalucía, Luis Montoto, 87, 4P, 41071 Seville, Spain.
Tel: +34 95 45 58 845 or 95 45 58 832. Fax: +34 95 45 58 853.
e-mail: briones@cica.es
Ignacio Marin, Hospital Universitario de Valme. (GRECA group), 41014 Seville, Spain.
Tel: +34 95 45 96 397 or 95 45 95 090. Fax: +34 95 46 93 757.
e-mail: miniml@valme.cica.es